# INTELLIGENT SKIN CANCER DETECTION USING ENHANCED PARTICLE SWARM OPTIMIZATION

**TECK YAN TAN**

**PhD**

**2019**

# INTELLIGENT SKIN CANCER DETECTION USING ENHANCED PARTICLE SWARM OPTIMIZATION

**TECK YAN TAN**

**A thesis submitted in partial fulfilment of the requirements of the**

**University of Northumbria at Newcastle for the degree of**

**Doctor of Philosophy**

**Research undertaken in the Faculty of Computer and Information Sciences**

**JUNE 2019**

# ABSTRACT

This research undertakes intelligent skin cancer diagnosis based on dermoscopy images using several variants of the Particle Swarm Optimization (PSO) algorithm for feature optimization. Since the identification of the most significant discriminative characteristics of the benign and malignant skin lesions plays an important role in robust skin cancer detection, the proposed PSO algorithms are employed for feature optimization. Specifically, the overall system contains multiple steps, i.e. pre-processing (noise removal), segmentation, feature extraction from both skin and lesion regions, proposed PSO based feature selection and classification. After extracting a large number of raw shapes, colour and texture features from the lesion areas, feature selection is conducted to identify the most discriminating significant feature subsets. Besides PSO and Genetic Algorithm (GA) based feature optimization, a total of four novel PSO variant algorithms, i.e. hybrid learning PSO (HLPSO), a PSO variant model (PSOVA), adaptive coefficient PSO (ACPSO), and random coefficient PSO (RCPSO), have been proposed for feature selection. Diverse search strategies are proposed in these models to mitigate premature convergence problems of the original PSO algorithm. Single and ensemble classifiers have been employed to perform benign and malignant lesion classification. Evaluated with multiple skin lesion and UCI databases and diverse unimodal and multimodal benchmark functions, the proposed PSO variants show superior performances over those of other advanced and classical search methods for identifying discriminative features that facilitate benign and malignant lesion classification as well as for solving diverse optimization problems with different landscapes. The Wilcoxon rank sum test is adopted to further ascertain superiority of the proposed algorithms over other methods statistically.

*To my wife Chloe, sisters Emily & Shan, two cats Jean & Xiao Mao, family and*

*Lastly,* ***Father - Separated by Death, together by Love****.*

# DECLARATION

I declare that the work contained in this thesis has not been submitted for any other award and that it is all my own work. I also confirm that this work fully acknowledges opinions, ideas and contributions from the work of others.

Any ethical clearance for the research presented in this thesis has been approved. Approval has been sought and granted by the Faculty Ethics Committee on 16/02/2014.

The Word Count of this dissertation is **48,823** Words.

Name: Teck Yan Tan

Date: 01/06/2019

# PUBLICATIONS

I declare that parts of the following papers have been included in this thesis.

**TAN, T.Y.**, ZHANG, L. AND JIANG, M., 2016, August. An intelligent decision support system for skin cancer detection from dermoscopy images. In Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), 2016 12th International Conference on (pp. 2194-2199). IEEE.

**TAN, T.Y.**, ZHANG, L., NEOH, S.C. AND LIM, C.P., 2018, June. Intelligent skin cancer detection using enhanced particle swarm optimisation. In Natural Computation, Knowledge-Based Systems. (Published)

**TAN, T. Y**., ZHANG, L., LIM, C. P., FIELDING, B., YU, Y. & ANDERSON, E. 2019. Evolving Ensemble Models for Image Segmentation Using Enhanced Particle Swarm Optimization. *IEEE Access***,** 1-1. (Published)

# ACKNOWLEDGEMENTS

First, I would like to thank my sponsorship, European Union sponsored (Erasmus Mundus), Centre of excellence for learning, innovation, networking and knowledge (cLINK) Programme and Professor Alamgir Hossain for giving me the opportunities to start my PhD study in University of Northumbria at Newcastle.

Secondly, I also would like to thank my Principal Supervisor Dr Li Zhang for supporting me and providing guidance for the whole four years. She has shown me her patience, motivation, and immense knowledge and led me to the truthful pathway and directed me in the research areas of artificial intelligence, computer vision and machine learning.

Finally, I appreciate my family for providing a psychological support whenever I feel difficult and giving me care and strength.

# CONTENT

# LIST OF FIGURES

# LIST OF TABLES

# GLOSSARY OF ASSOCIATED TERMS

| Term | Description |
|---|---|
| **ABCD** | Asymmetry, Border, Color, Diameter |
| **ABC** | Artificial Bee Colony |
| **ACPSO** | Adaptive Coefficient Particle Swarm Optimisation |
| **ADOLPSO** | Adaptive Division of Labour Particle Swarm Optimisation |
| **ACO** | Ant Colony Optimisation |
| **AGPSO** | Autonomous Group Particle Swarm Optimisation |
| **BPSO** | Bare Bones Particle Swarm Optimisation |
| **BBPSOV** | Bare Bones Particle Swarm Optimisation Variant |
| **BA** | Bat Algorithm |
| **RCPSO** | Random Coefficient Particle Swarm Optimisation |
| **CBPSO** | Chaotic Binary Particle Swarm Optimisation |
| **CLPSO** | Comprehensive Learning Particle Swarm Optimisation |
| **CNN** | Convolutional Neural Network |
| **CAD** | Computer-Aided Diagnosis |
| **CA** | Cultural Algorithm |
| **DT** | Decision Tree |
| **DNA** | Deoxyribose Nucleic Acid |
| **DE** | Differential Evolution |
| **DA** | Dragonfly Algorithm |
| **DNLPSO** | Dynamic Neighbourhood Learning Particle Swarm Optimizer |
| **ELPSO** | Enhanced Leader Particle Swarm Optimizer |
| **ELM** | Epiluminescence Microscopy |
| **FPA** | Flower Pollination Algorithm |
| **FCRN** | Fully Convolutional Residual Network |
| **GA** | Genetic Algorithm |
| **GP** | Genetic Programming |
| **GM** | Geometric Mean |
| **GWO** | Grey Wolf Optimizer |
| **GLCM** | Grey-Level Co-Occurrence Matrix |
| **GLRLM** | Grey-Level Run Length Matrix |
| **HS** | Harmony Search |
| **HEPSO** | High Exploration Particle Swarm Optimisation |
| **HOL** | Histogram of Oriented Lines |

| | |
|---|---|
| **HOG** | Histogram-oriented Gradients |
| **PH2** | Hospital Pedro Hispano |
| **HLPSO** | Hybrid Learning Particle Swarm Optimisation |
| **KNN** | K-Nearest Neighbour |
| **LBP** | Local Binary Pattern |
| **MSKs** | Melanocytic Skin Lesions |
| **MFO** | Moth-Flame Optimization |
| **NN** | Neural Network |
| **NoMSLs** | Nonmelanocytic Skin Lesions |
| **PSO** | Particle Swarm Optimisation |
| **PSOVA** | Particle Swarm Optimisation Variants |
| **RBF** | Radial Basis Function |
| **RF** | Random Forest |
| **SK** | Seborrheic Keratosis |
| **STD** | Standard Deviation |
| **SVM** | Support Vector Machine |
| **ThBPSO** | Threshold Binary Particle Swarm Optimizer |
| **UV** | Ultraviolet |

# CHAPTER 1
# INTRODUCTION

## 1.1 Background

Cancer is among the leading cause of death worldwide. Skin cancer has been one of the most common diseases with more than three million cases reported each year globally (Siegel et al., 2016). Despite the decrease in the occurrences of common cancers, incidences of melanomas, the most dangerous form of skin cancer, have been increasing at a higher rate than those of other five common cancers: breast, lung, non-Hodgkin lymphoma, bowel, and prostate cancers (Linos et al., 2009). As one of the top listed cancers in United Kingdom, United States and Australia, skin cancer is ranked as one of the deadliest cancers in the region (Diepgen and Mahler, 2002, Cancer et al., 2014). The latest statistic shown by Cancer Research stated that the forms of skin cancer such as malignant melanoma are causing 2,459 deaths in the United Kingdom (UK) every year (Berry, 2016). In the UK, 46 of 100 people with melanoma are diagnosed with a late stage of skin cancer and it is too late to be treated. This amount can be petrifying if it is combined with all affected countries. For example, this has led to a total of 232,000 cases with 55,000 deaths in New Zealand and Australia in recent years (Cancer et al., 2014, Campbell, 2014, Garg et al., 2018).

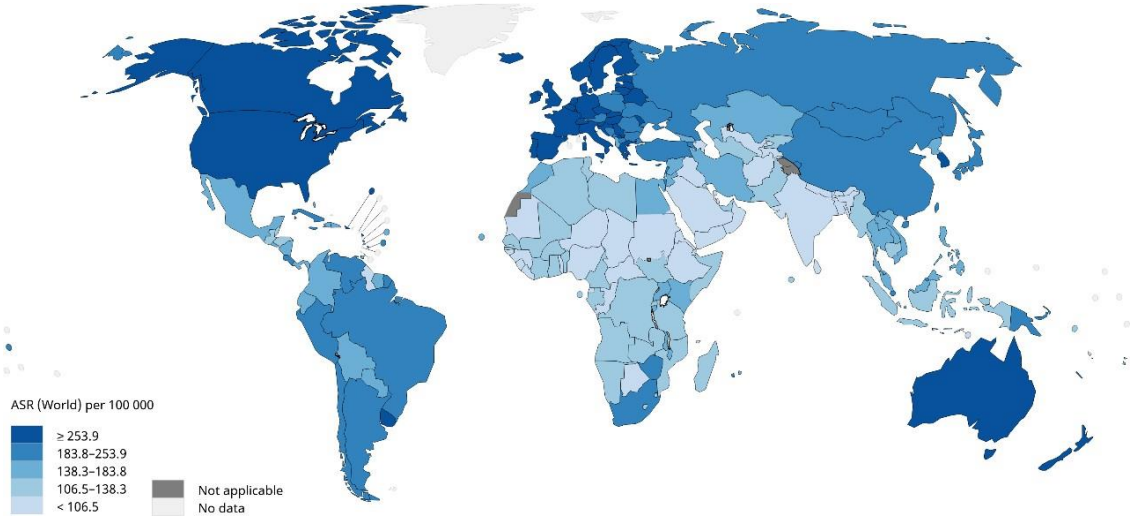



**Figure 1-1: The distribution of melanoma and other skin cancers (with red indicating the most affected areas and yellow indicating the lease affected areas) (Bray F, 2018)**

As shown in Figure 1-1, the places that are most affected by skin cancer diseases has been illustrated in red, such as the United States, Australia, and some European countries. As observed by Valentine (2017), 74,000 Americans were diagnosed with skin cancer in 2015. The Australian Institute of Health also published a study in the same year which confirmed that the rate of skin cancer is the highest amongst Australians, reporting upwards of 100,000 skin cancer cases in one year, and Australians also were among the highest averages for cancer diagnosis in the world (Morris et al., 2009, Aitken et al., 2018, Bray F, 2018). Heckel et al. (2017) determined that by the end of 2017, approximately 135,000 Australians would be diagnosed with skin cancer every year, and by the end of 2020, this number may increase to 150,000. This statistic indicates that Australians are more at risk when compared to Americans, who have approximately 87,000 skin cancer cases by 2017. A recent study determined that increased exposure to ultraviolet (UV) from direct sunlight or tanning beds will result in a 45% increase in melanoma (Le Clair and Cockburn, 2016). Although people all over the world are reporting higher occurrences of skin cancer, Australians are reporting a significantly higher number of cases (Ferlay et al., 2010, Jemal et al., 2011). This is because of a hole in the ozone layer over Antarctica that has allowed higher amounts of UV radiation. Siegel et al. (2018) observed that despite being 7,246km away, Australia is affected by the harmful effects of these rays.

There are many types of skin lesion conditions, such as dermatofibroma, pyogenic granuloma, granuloma telangiectatic, seborrhoea actinic keratosis, and solar keratoses. Melanoma is a lot less common than other skin cancers, but it is a lot more dangerous. Approximately three out of four deaths due to skin cancer are related to melanoma (Diepgen and Mahler, 2002). However, any skin cancer category such as the deadliest melanoma is potentially curable with an early diagnosis (Jerant et al., 2000, Hoffmann et al., 2003, Smaoui and Bessassi, 2013, Alfed et al., 2015). Nevertheless, a well-trained dermatologist clinical consultation could be unaffordable for some places (BURGISS et al., 1997, Feng et al., 2016).

## 1.2 Motivation

Access to efficient medical assistance has been an issue for rural areas. The medical healthcare services could be costly in some places to provide fast access or private treatment. For example, for the developing nation, India, although the probabilities of getting skin cancer in India are less than some Western countries, the cases of sunburn and skin cancer still do occur. The ratio between the patient and doctor is significantly imbalance in India, e.g. for one doctor, it might have one thousands of patients to look after (Chandramohan, 2013, Deo, 2013). Thus, to make doctor examination prioritise is essential. Therefore, a reliable, fast, effective, prediction suggestion application for an earlier diagnosis could be developed for the individual users or to assist doctors. For instance, this can also detect early signs of skin cancer and suggests the patients to go and visit a doctor sooner.

Image recognition has drawn more attention and become one of the most popular research topics in computer vision and artificial intelligence fields (Chiem et al., 2007, Abuzaghleh et al., 2014a, Sáez et al., 2014, Gutman et al., 2016, Satheesha et al., 2017, Katapadi et al., 2017, Taufiq et al., 2017, Abbott and Smith, 2018). However, it is a challenging task to recognise multiple types of skin lesions with numerous of skin features extracted based on the ABCD (Asymmetry, Border, Colour, Diameter) guidelines, which are widely used for clinical diagnosis and signify the appearances of different types of lesions (Gola Isasi et al., 2011, Mahmoud et al., 2018). Specifically, multiple dimensions of features such as

colour, size and texture are extracted from skin and lesion areas, and these features could be redundant or even contradictory to the classification process, which may confuse classifiers and lead to inaccurate conclusions.

Recently, many researchers have adopted multiple methods for early melanoma diagnosis. Some studies indicated that the classification performances can be improved by removing redundant or irrelevant features from the originally extracted feature sets for melanoma detection (Andrushia and Patricia, 2018, Tan et al., 2018, Milton, 2019). For this reason, in order to identify the most important characteristics for lesion classification, evolutionary algorithm based optimization methods are explored owing to their capabilities in dealing with performance, speed and consistency of results (Zhang et al., 2015c, Mistry et al., 2017, Srisukkham et al., 2017, Zhang et al., 2018a). The purpose of feature selection is mainly attempting to remove unnecessary features without decreasing or compromising the prediction accuracy and identify a feature subset within the initial collection of the features. Since feature selection algorithms could enhance prediction accuracy, and reduce the computation cost for classification, as a result, such research studies have recently developed into a new industry. The explorative methods of feature subsets indicate that the current feature selection methods can be divided into the two categories, i.e. the filter and wrapper methods (Suto et al., 2016). The wrapper method utilises a classification algorithm to assess the feature subsets and employs an exploration method for identifying the ideal subsets. As this method examines a classifier with an assessment or exploration procedure, the wrapper-based method typically obtains enhanced outcomes compared to the filter method. On the other hand, the filter method primarily depends on the features of the datasets for assessing and identifying feature subsets in the absence of contemplating a unique learning approach.

## 1.3 Research Aims and Objectives

The purpose of this research is to find an appropriate methodology for the development of prediction and recognition of melanoma from dermoscopy images. It focuses on evolutionary algorithm-based feature selection to optimize the initial feature sets extracted using the ABCD rules. The research aims, and objectives are shown as follows:

- To study the basic biologist knowledge regarding the skin cancer and relevant existing literatures on skin cancer classification.
- To examine necessary development tools and methods and review literatures for skin lesion classification.
- To investigate and develop machine learning and feature selection methods for skin cancer detection.
- To develop and implement an efficient, intelligent scheme to address the existing challenges (such as premature convergence in feature selection and performance improvement).
- To compare the outcomes of the proposed algorithms with existing state-of-the-art methods.

## 1.4 Research Contributions

The main contributions of this research have been categorised into five aspects as follows:

**A.** The first contribution of this research is to use Particle Swarm Optimization (PSO) and Genetic Algorithm (GA) for feature selection to enhance accuracy for skin lesion classification.

**B.** The second contribution is the proposal of a modified PSO model, known as hybrid learning PSO (HLPSO), for feature selection. In order to overcome premature convergence of the original PSO model, HLPSO integrates PSO with multiple strategies, i.e. sub-swarm division, random explorations using Gaussian, Cauchy and Levy distributions, crossover and mutation operators and scattering strategies.

Instead of using the whole swarm for exploration, HLPSO has divided the original population into two sub-swarms. The search process in one sub-swarm is guided by long and short exploration using probability distributions, while the other sub-swarm is led by the original PSO. Moreover, the 10 best particles of the overall

swarm will be selected for further crossover and mutation to create 20 offspring solutions. A scattering mechanism is also used to re-locate the worst particles in the swarm to diversify the population.

The proposed HLPSO model outperforms classical methods including Bat Algorithm (BA), Dragonfly Algorithm (DA), Harmony Search (HS), Flower Pollination Algorithm (FPA), Moth-Flame Optimisation (MFO), Artificial Bee Colony (ABC), Cultural Algorithm (CA), GA and PSO, significantly, for solving feature selection and benchmark functions.

**C.** The third contribution of the present study is to put forth another PSO variant model, i.e. the PSOVA algorithm, for feature selection and skin cancer classification.

To prevent the premature convergence of the conventional PSO algorithm, PSOVA includes more diverse velocity updating strategies than those of HLPSO for feature optimization. Specifically, PSOVA combines the following essential steps to diversify the search, i.e. sub-swarm division, attraction and evading actions guided by multiple swarm leaders and worst experiense, swarm leader enhancement and diverse matrix representations. In comparison with the classical methods employed for the evaluation of HLPSO and state-of-the-art PSO variants, the PSOVA model is more effective in identifying the most significant features for detecting both benign and cancerous lesions. With an emphasis on the selection of features based on attraction and evading action, the key mechanisms of PSOVA are discussed below.

To decrease the chances of trapping in local optima, two remote leaders are identified initially which have similar fitness scores but low correlation in positions. In other words, the selected two swarm leaders must assure to have an appropriate distance between each other, and so the search processes could reach different regions. The two remote leaders are then used to guide the sub-swarm based searches.

To diversify the search processes, PSOVA enables the particles to follow local and global best solutions and worst signals in each dimension and randomly selected partial dimensions. The best solution identified in each sub-swarm is also further improved, and the exploitation is augmented by employing the three random walks of Gaussian, Cauchy and Levy distributions.

The search is further diversified using a dynamic matrix representation of the swarm in each generation. In contrast to other search strategies and methods, the proposed PSOVA algorithm is highly efficient not only in choosing ideal features for categorisation of melanomas, but also in addressing unimodal and multimodal benchmark functions. Additionally, as compared with the approaches used in earlier studies on skin cancer classification, the algorithm has achieved the highest performances when it comes to detecting skin cancer.

**D.** The fourth major contribution of this research is the extraction of diverse lesion features using multiple feature descriptors and the proposal of an adaptive coefficient PSO model, known as ACPSO, for feature selection. An ensemble classification model is also developed by integrating several base classifiers dedicated to each feature type to enhance performance.

First, the clinical diagnosis prompts the initial extraction of a series of texture features (e.g. GLRLM) and the ABCD features, including shape and colour features. To better represent the lesions, other more refined textural features are also extracted using Histogram-oriented Gradients (HOG) and Local Binary Pattern (LBP) descriptors, respectively.

A new improved PSO algorithm (ACPSO) is also proposed for feature selection. In comparison with HLPSO and PSOVA, ACPSO employs adaptive acceleration coefficients to improve the skin lesion detection performance. Besides integrating features for a global search using adaptive coefficients, the ACPSO model also divides features into sub-dimensions to enhance feature optimization for each

lesion region (e.g. top, middle and bottom regions of the lesion). The search process of ACPSO begins with the identification of three remote swarm leaders with similar fitness scores but low correlation in positions to guide three sub-swarm-based searches, with each sub-swarm led by one swarm leader.

The ascending and descending acceleration coefficients are proposed in ACPSO. These adaptive coefficients are generated dynamically based on partial circle, sine and helix waveforms to guide sub-swarm-based search. These coefficients adaptation processes ensure the search will focus on global exploration from the start of the iterations and shift to local exploitation in later stage of the iterations to accelerate convergence.

Furthermore, a sub-dimension-based search is also performed to obtain more refined important features to improve lesion classification. Finally, a few new positions are re-initialised randomly to replace the weak solutions in the population. ACPSO outperforms the previously employed classical search methods and advanced PSO variants for diverse feature selection and unimodal and multimodal optimization problems.

**E.** The fifth major contribution of this study is to further improve the above proposed ACPSO model by using dynamically generated random coefficients. This improved model is named as random coefficient PSO (RCPSO). Specifically, RCPSO employs dynamic random acceleration coefficients generated based on the full circle, sine and helix waveforms instead of ascending and descending coefficients produced by partial waveforms as in ACPSO.

The RCPSO model uses positive and negative random acceleration coefficients supported by utilising full waveforms resulting from non-linear circle, sine and helix functions to explore the search space. Since the random coefficients generated in each iteration are more diverse including both positive and negative values, RCPSO explores a wider search space, therefore having more chances of finding global optima.

In addition to making it possible to explore the search space more widely by enhancing diversification, random coefficients also fine-tune the areas of local and global best solutions. Hence, the RCPSO model is better equipped to find the global optima because it replicates the hovering flight behaviour of hummingbirds around a target (e.g. flower or food source) and can investigate a broader search space.

## 1.5 Evaluation Resources

By default, MATLAB utilised only a single core computation power to complete each task. This could be a problem when performing large-scale computational tasks, since it could be very expensive in terms of time and cost. A total of 30 runs were performed by each method to identify the most important feature subsets in our experiments. Therefore, a MATLAB parallel computing toolbox was used to increase the efficiency and reduce the experiment time. Due to the experimental devices improved from chapter to chapter, multiple computer devices have been utilised. These include a multicore workstation with i7 quad-core CPUs with 16GB RAM and another workstation with 36 core Intel Xeon processors and 256GB RAM. By doing this, it can spread a 30-run experiment with 10-fold cross-validation into multiple treats.

## 1.6 Datasets for Evaluation

The experiments are conducted using three publicly available skin lesion datasets i.e. Hospital Pedro Hispano (PH2) (Mendoncÿa et al., 2013), Edinburgh Research and Innovation (Dermofit) dataset (Ballerini et al., 2013) and Dermnet (Dermnet, 2016). Two additional medical datasets, i.e. breast cancer and epileptic seizure, from the UCI machine learning repository, and one non-medical UCI dataset, i.e. spam base (Lichman, 2013), are also used for further evaluating the performance of feature optimisation for the proposed PSO algorithm.

## 1.7 Thesis Outlines

This thesis is organised as follows.

**Chapter 2** explores a brief overview of related research on skin cancer and lesion classification and diverse optimization and feature selection methods. Specifically, it provides literature review on related work discussion on feature extraction, feature selection and skin lesion classification.

**Chapter 3** presents the first contribution of this research, i.e. the first proposed skin cancer detection system, including the extraction of ABCD features, GA and PSO based feature selection and SVM-based classification.

**Chapter 4** presents the second contribution of this research, i.e. the HLPSO based feature selection. HLPSO incorporates PSO with sub-swarm division, probability distributions, crossover, mutation and scattering strategies for feature selection. An evaluation with comparison is conducted to verify the algorithm's performance.

**Chapter 5** presents the third contribution of this research, i.e. the proposed PSOVA-based feature optimisation to address the classification of skin lesion problems. Several distinctive velocities updating strategies were proposed to diversify the sub-swarm based search processes. Dynamic matrix representation is also used to overcome stagnation.

**Chapter 6** presents the fourth contribution of this research by using two extra LBP and HOG feature descriptors and the proposal of ACPSO-based feature selection, as well as ensemble-based classification. The ACPSO model comprises both a global search based on increasing and decreasing acceleration coefficients and a sub-dimension based search for lesion classification.

**Chapter 7** presents the final contribution of this research, i.e. the proposed RCPSO based features selection algorithm, where random acceleration coefficients are generated using the full waveforms of non-linear circle, sine and helix functions to guide sub-swarm-based searches.

Finally, **Chapter 8** provides conclusions, summarizes research contributions and identifies directions for future work.

# CHAPTER 2
# LITERATURE REVIEW

## 2.1 Introduction

In this chapter, the research examines previous research on skin lesion classification and feature selection techniques to determine related methodologies for melanoma classification. Analysis of these previous works will also take place to gain a better understanding of these related techniques as well as create a background theory on this subject so that a new, relevant technique can be devised.

## 2.2 Biology of Skin Cancer

### 2.2.1 Cause of Skin Cancer

The sunlight lights up our world, feeds the plants, helps body to create vitamin D, and does a ton of other amazing things (Holick et al., 2011, Webb et al., 2018). But in the meanwhile, it may also cause sunburns, ageing, saggy skin and many types of skin cancers. Several wavelengths or energies are released through the sunlight. The light that reaches us on the earth constitutes 4% ultraviolet light, 43% visible light, and 53% infrared light. The sunlight is made up of seven colours, i.e. red, orange, yellow, green, blue, indigo and violet. And six in those colours are visible to the eye, but only violet is not. The ultraviolet light can be broken up into three groups based on their wavelengths, i.e. near (UV-A), mid (UV-B) and far (UV-C). UV-C is the bounciest, and it can be blocked by the atmosphere and glass effortlessly. UV-B is partially absorbed by our atmosphere especially the ozone layer, but this may cause sunburns and skin cancer. Then

there is UV-A, which is the type that cannot be blocked by the glass, therefore, glasses cannot be reliable for protection. The sunblock will be needed in this case, otherwise this may be most likely to cause skin cancer (Diffey and Farr, 1991, Pfeifer et al., 2005).

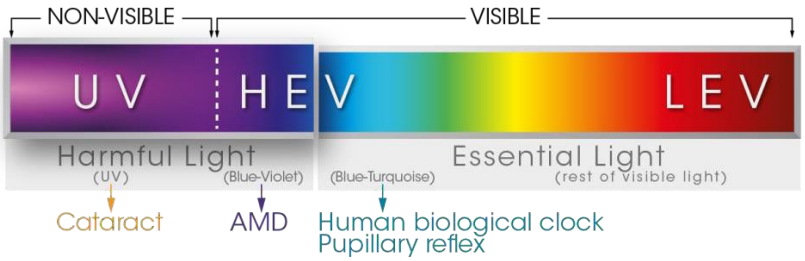



**Figure 2-1: Relationship between light & health (1) (Exposed, 2018)**

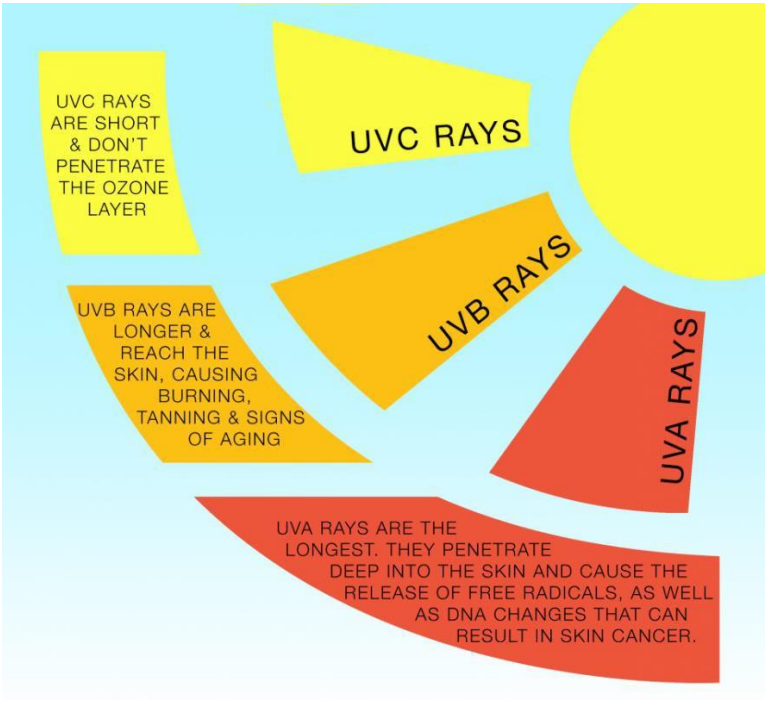



**Figure 2-2: Relationship between light & health (2) (Nevada, 2017)**

The human skin is an incredible organ, and it does the best to protect us from the sunlight with a protein called melanin. Melanin is in the skin to absorb both UV-B and UV-A lights, thus causing a lot of the protein that makes the skin looks darker. The genetics that define how much melanin each person has are standard in the body. Ultraviolet light could penetrate the skin and cause deoxyribose nucleic acid (DNA) damage (Karran and Brem, 2016). When the UV-B light hits the skin, although it only goes through the upper layer,

it causes direct damage to the DNA. So, the body protects itself by producing more melanin, although it can take a few days for the melanin to be produced and make the skin look tanner.

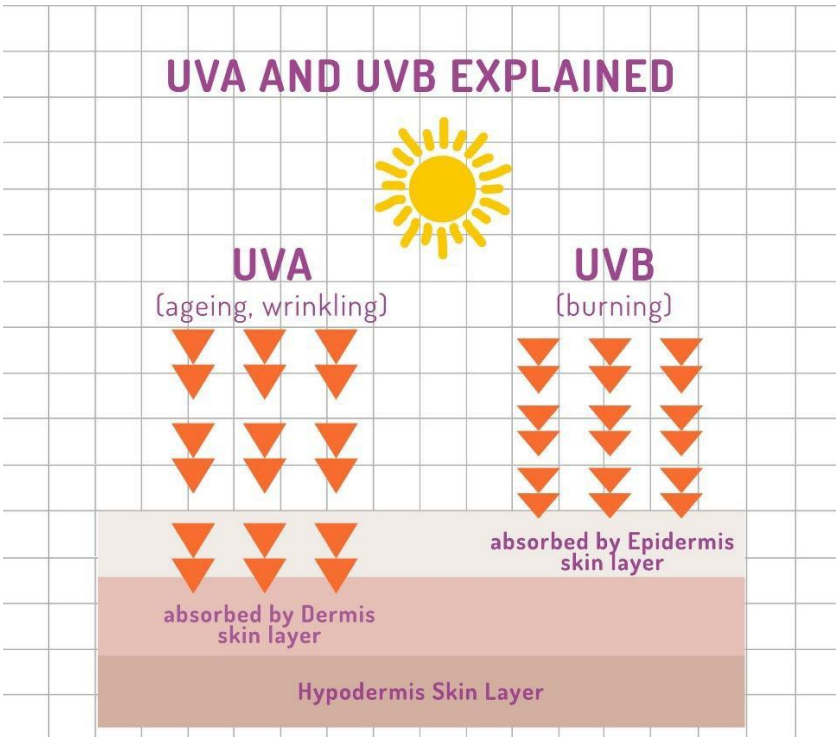



**Figure 2-3: Relationship between light & health (3) (Care, 2018)**

Nonetheless, the UV-A light is slightly different from UV-B and UV-C, since it has the higher energy to pass through the atmosphere and can penetrate the skin even deeper. The skin releases stored melanin when the UV-A light hits the skin, and UV-A causes formation of reactive oxygen which in turn causes DNA damage (Brash et al., 1991, de Gruijl et al., 2001, Kimeswenger et al., 2018, Khan et al., 2018). The human body has an intelligent mechanism that detects UV damage when something is wrong and creates warning molecules. When single cells in our body divide enzymes that create the DNA, mistakes are often made on which may be beneficial or harmful. Although any DNA damage can cause cancer, several defences in the human body work to repair the UV damage (Courdavault et al., 2005, Nichols and Katiyar, 2010). The body will initiate an inflammatory response if several cells are sending out a distress signal. In this response, various blood cells are sent to the damaged area to repair it, causing an influx of blood

that is seen as the sunburn or redness on the skin, depending on the skin colour. If these safeguards are surpassed by a considerable number of cancer cells and the stem cells cannot fix the damage, then it results in skin cancer. The cells, once damaged, can begin reproducing and mutating at the damaged location, which then leads to the cell becoming cancerous (Friedberg et al., 2005).

### 2.2.2 The Importance of Screening Skin Cancer Earlier

As deadly type of skin cancer, melanoma, has a higher incidence with an increasing number of people affected by it (Queen, 2017). Over time the number of people who will develop melanoma at some point in their life is estimated to be about one in 36 men and one in 55 women (Chen et al., 2018). People who are screening with a melanoma in the earliest stage have a higher probability of surviving it. The figure below shows the reason and the importance of screening skin cancer in an early stage.

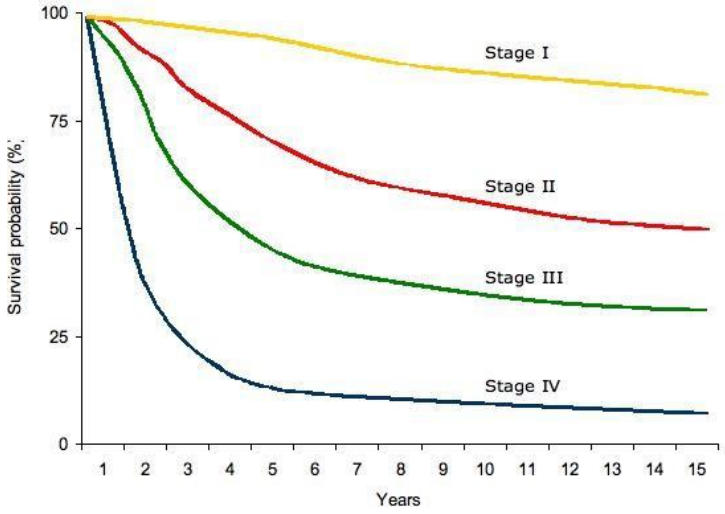



**Figure 2-4: Skin cancer survival probability (Travel, 2015)**

Figure 2-4 shows, what the dermatologist called, the survival curve or survival probability. Skin cancer can be divided into four stages, and more severe stages can be more difficult to be treated. Each colour of the line in Figure 2-4 presents a different stage of skin cancer; it is also shown how early people could find skin cancer. So, the first stage of melanoma can be cleared easily. It is the earliest form, which is grown only on the top

layer of the skin. If found at the earliest, a 100% cure rate can be achieved. Some melanomas start a bit later, e.g. in 10 years' time or later in the stage, Therefore, if the melanoma is detected in the earliest period, the chance of removing, curing and surviving will be increased significantly (Ferlay et al., 2010, Siegel et al., 2018).

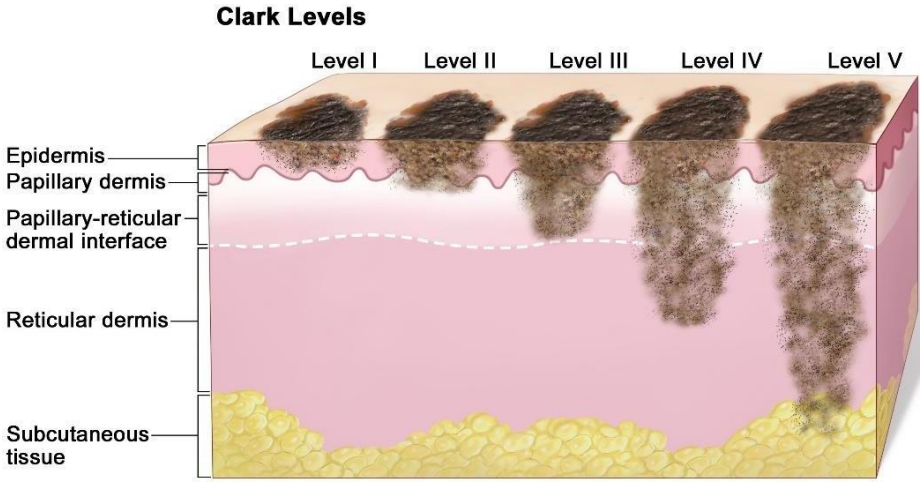



**Figure 2-5: Skin Clark Levels (WInslow, 2009)**

The human skin is a form of three layers, i.e. Epidermis, Dermis and Subcutaneous Tissue (Griffiths et al., 2016). Figure 2-5 shows different stages of melanoma. As we can see, the tumour starts as a brown confined lesion in the top layer of the skin. Over time, it starts getting deeper and deeper into the skin beyond the top layer of the epidermis and grows into the reticular dermis. Ultimately, the tumour cells will go deeper into the skin, and invade into the lymphatics or grow into the clod vessels in the skin. When it reaches the stage four, the melanoma may have started to spread beyond the skin, first to the lymph nodes and then it can be spread all over the body such as brain, liver and lungs (Habif, 2015). When it comes to stage four, this is the most progressive melanoma, and this is where the chance of survival becomes lower (Patterson, 2014).

### 2.2.3 Diagnosis of Skin Cancer

Skin cancer can be diagnosed in several ways. First, most skin lesions are visible to most people, and the patient can do a self-examination by looking around the body to find anything abnormal by using a mirror or someone who could help before consulting a doctor. The dermatologist recommends doing this checking regularly, so that if an evolving lesion can be found early. Multiple studies found that most melanoma skin cancer is actually first detected by the patient (Brady et al., 2000, McGuire et al., 2011), and studies shown (Titus et al., 2013) that skin self-examinations including checking lesion size and depth of melanoma can increase the surviving rate. Turn out, that the people who did the self-examinations 1 to 11 times per year had a statistically significant low chances of forming melanomas (Titus et al., 2013). Therefore, this is good evidence that skin self-examination works.

The self-examination could be easily conducted by following the guidelines of ABCDE rules provided by dermatologists (Glazer et al., 2017). All dermatologists are also using these steps to diagnose different types of skin lesions. First, A in the ABCDE rules represents asymmetry, with the two halves of the skin lesion that do not look like each other. If a line drew horizontal and vertical through the middle of the lesion, the halves of the lesion would not match in size. Then this indicates asymmetry. Furthermore, B indicates the border; this refers to the edges of irregularity and C denotes colour. Cancerous lesions tend to possess a variety of colours especially blue or white. Moreover, D refers to diameter. This is involved when the lesion diameter is larger than 6mm, i.e. ¼ inch or larger than a pencil eraser. However, the diameter factor could be misleading sometimes. Some melanomas can be smaller than 6mm but with irregular shapes. Finally, E indicates evolving since the size of the lesion or mole can be changing over time. If the lesion is getting larger than the initial appearance, then this lesion must be evaluated further by a dermatologist (Griffiths et al., 2016).


**Figure 2-6: The ABCDE Diagnosis Patterns for Skin Cancer**

Some dermatologists are still using the old fashion gadget for diagnosis which what they called a "wood lamp". It consists of a big magnifying glass and black light to show the pigmentation on the skin (Lee et al., 2018). For some more advanced dermatologists, they have their own new gadget called the dermatoscopy, and it uses a polarised small handheld device to look up closely at the different types of pigmentation of a mole. This helps a lot for a dermatologist to decide whether the patients will go for a skin biopsy (Limone and Meadows, 2017). A skin biopsy can be performed with several types, i.e. shave, excisional or punch. A dermatologist injects lidocaine to numb the area instantly, and the dermatologist either chooses to 1) shave off the lesion by cutting parallel to the skin, 2) punch biopsy by using a mosquito-like puncher tool punching into the skin for full thickness checking, 3) lastly, cut out an ellipse of skin around the skin lesion using a scalpel and close it with stitches for excisional biopsy (Milan, 2017, Yagerman and Stevenson, 2018).

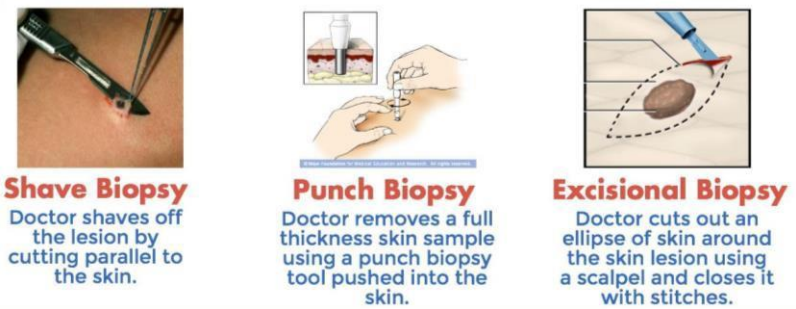



**Figure 2-7: Skin sample extraction method (Pickett, 2011)**

Therefore, a small skin sample will be taken from a suspicious spot and sent to a histopathology laboratory where the doctors examine it under the microscope and look at the architecture of the cell and decide whether is abnormal or normal. This process takes two days to up to two weeks where they need to look at the special stains. If it is skin cancer, there are different ways to remove it. Usually, if it is on an essential area of the body, and it is squamous cell cancer or basal cell cancer, then the dermatologist will perform a surgery called Mohs surgery. There are over 20 types of skin lesions, but only six types are the most popular ones and could exist anywhere. The listed below are the popular types of skin lesions (Griffiths et al., 2016, Glazer et al., 2017).

**Table 2.1: Top six types of skin diseases**

| | |
|---|---|
|  | **Atypical Moles**<br>A benign growth resembles melanoma. Individuals with such atypical moles have higher possibilities of developing melanoma. |
|  | **Actinic keratosis**<br>This forms in areas that are more exposed to the sun, including the face, lips, back of the hands, and scalp. It is a precancerous growth that is scaly and crusty. |
|  | **Merkel cell carcinoma**<br>This is a form of skin cancer that is rare and aggressive and can spread or recur at a higher risk. |
|  | **Squamous cell carcinoma**<br>This type of lesions ranks the second among the most common types of skin cancers. It may appear anywhere on the body but is typically seen in parts that receive more sun exposure. It is fatal if left untreated. |

**Basal cell carcinoma**
This is the most common form of skin cancer whose growths are seen on the skin's outermost layer and that seldom spreads to other areas on the body.

**Melanoma**
Melanoma is the fatal type of skin cancer. It can result from a genetic predisposition or more typical exposure to UV rays.

### 2.2.4 Sign and Symptoms

It is challenging to notice a sign and symptoms for skin cancer. Mostly the skin will start turning into a different colour and create a small lesion which is not visible and very hard to be noticed by the human eye. As soon as the skin starts to change colour (e.g. dark or multiple colours), forming any strange shapes (e.g. irregular borders), any inflamed skin that will not heal over time, a medical diagnosis can be performed but may not be early enough (Griffiths et al., 2016, Glazer et al., 2017). Therefore, accurate instant and early diagnosis are vital.

## 2.3 Existing Work on Automatic Skin Cancer Identification

The effective treatment of malignant melanoma skin cancer and survival depend on accurate early diagnosis. Researchers have proposed a range of computerised methods for diagnosis and classification, which performed acceptably regarding detecting skin cancer. Despite this, minimisation of the rate of incorrect detection remains a problem, especially as false positives prompt additional investigation through intervention by a qualified pathologist. Since 1985, there were a lot of skin medical imaging and vision classification research works conducted. Friedman and Rigel (1985) performed skin lesion classification using feature selection and successfully selected five features from initial 87 features obtained through the surface profiles of skin lesions. As one of the most well-known studies through the computation of correlation coefficients, Green et al. (1994) revealed that lesion size is a critical feature in the context of their framework.

Schmid (1999) has utilised multiple methods to detect lesion borders and analysed the dimension of symmetry in a skin lesion pigmentation computer-aided diagnosis (CAD) system. In their work, a lesion image was first processed by rendering the outer layer of skin translucency using Epiluminescence microscopy (ELM), also known as magnified oil immersion diascopy. Then the lesion was detected either using image segmentation or a border detection method. Once the lesion was located, and its features were extracted, which were then used for classification. Gilmore et al. (2010) have developed a system using image processing and SVM for the classification of melanoma. Their work achieved a specificity of 72% and a sensitivity of 86% after an examination of 199 skin lesion images. Ramlakhan and Shang (2011) developed a portable classifying scheme and application for epidermal lesion classification. Their work was composed of image segmentation, feature extraction and classification. Empirical testing of this system presented an accuracy of 67% for malignant classification.

Doukas et al. (2012) conducted automatic skin lesion assessment based on cloud and smartphone platforms. Their work employed the Active Shape Model to extract texture, shape, and size features. A set of classifiers, including Neural Network (NN), Support Vector Machine (SVM), Decision Tree (DT), was used to conduct the classification of different lesion types such as melanoma, dysplastic and common (benign) nevus. There are also many other developments on automatic skin cancer diagnosis in the literature, whereas Glaister et al. (2013), (2014) proposed a multistage illumination modelling algorithm for illumination correction and a texture-based skin lesion segmentation model, respectively.

Barata et al. (2014) presented alternative processes for melanoma identification via dermoscopy imagery, based on surface texture and shade of colour feature extraction. Texture and colour features were compared in performance in their work, with the latter proving more useful for skin cancer classification. It should note that either method nevertheless attained strong results regarding classification accuracy. The drawbacks of these systems were the inability to operate on a real-time diagnosis application owing to the time-consuming complex training process.

Kruk et al. (2015) used comprehensive feature extraction of ABCD, image filtering, feature selection and benign/melanoma classification. Their work used substantial descriptors, such as Kolmogorov–Smirnov statistical distance, fractal texture analysis, maximum sub-region principle, and percolation theory, to extract features. It also used Fisher discriminant measure, fast correlation-based filter, and correlation feature selection to select distinctive features. Random forest (RF) and SVM were also used to classify melanoma. Their results showed the best performance was achieved by combining SVM with the Fisher discriminant measure. Adjed et al. (2015) used a combination of textural and structural features to classify melanoma. They used wavelet and curvelet transforms for extracting structural feature and various LBP variants for extracting textural feature. This study was tested using an SVM model and a random sampling cross-validation strategy, and the system achieved an accuracy rate of 86.07% for the evaluation of the PH2 dataset. Barata et al. (2015) classified melanoma by employing both local and global methods and suggested that melanoma can be easily identified by using colour features in comparison to texture features.

Shimizu et al. (2015) developed a digital diagnosis system for the classification of four skin lesion conditions including melanoma, nevus, basal cell carcinoma (BCC), and seborrheic keratosis (SK). In their work, lesions were categorized into two general categories, i.e. melanocytic skin lesions (MSLs) and nonmelanocytic skin lesions (NoMSLs). Melanoma and nevus belong to MSLs while BCC and SK belong to NoMSLs. The detection of NoMSLs has been rarely addressed in the literature, despite their high occurrences. Their work was therefore dedicated to the identification of both MSLs and NoMSLs. It extracted a total of 828 features representing colour, sub-region, and texture information. A layered classification method was used to conduct the four-class lesion classification. Their classification model firstly categorized MSL and NoMSLs using a binary classifier. Another two binary classifiers were subsequently used to distinguish melanoma from nevus and BCC from SK, respectively.

Abuzaghleh et al. (2015a) proposed an intelligent system for early detection and prevention of malignant lesions such as melanoma. Their method consisted of a

prevention component and a detection component. The former integrated a novel equation to identity the risk of skin burns and generated alerts when necessary, while the latter comprised several key stages for the classification of different lesion types, including noise removal, lesion segmentation, feature extraction and classification. Evaluated with the PH2 Dermoscopy image database from Pedro Hispano Hospital, the proposed method showed impressive performance for the classification of benign, atypical, and melanoma cases.

Tan et al. (2016) developed an intelligent decision support system for melanoma detection. Their work consisted of several key stages and extracted a high-dimensional feature vector integrating shape, colour, and texture information owing to their high correlation with clinical characteristics associated with skin cancer identification. The Genetic Algorithm (GA) was subsequently used to identify the most contributing factors for diagnosis. The work showed performance improvements in comparison with those from other related work reported in the literature.

Kasmi and Mokrani (2016) detected melanoma by extracting ABCD features based on the dermoscopy ABCD rule for clinical diagnosis. They used Gabor and median filters for hair removal, and geodesic active contours for determining the lesion boundaries. The extracted features of the ABCD rule included brightness, shape, asymmetries, colour, the border, blue-white veil, geometrical properties, and pigment network. After that, they calculated a total dermoscopy score for classifying different types of lesions and determined that low scores indicated benign cases, middle scores signalling suspicious cases, and high scores indicating malignant cases.

Alfed and Khelifi (2017) proposed a feature-based approach based on several codebooks with effective characterisation of skin cancer lesions. Their work extracted colour and texture features, by using the histogram of oriented lines (HOL), HOG, third order Zernike moments, and colour vector angles. They claimed that the differences between categories were reduced and, the discrimination capability of the features in question was weakened by the orientation data associated with the traditional HOL and HOG operators. Therefore, each of the HOL and HOG histograms constituted the basis for a codebook,

and a third codebook was based on a combination of colour vector angles and third order Zernike moments for performance enhancement.

Xie et al. (2017) attempted the classification of melanoma using a neural network ensemble classifier with dermoscopy images. In their study, they first used a self-generating neural network to divide the lesions. Border irregularities were extracted from complete as well as incomplete lesions using this feature extraction process. The proposed meta-ensemble classifier merged three ensemble models with different base model types and various network topologies. The first ensemble model comprised a set of networks containing the same structure and type. On the other hand, the second ensemble model was built using diverse neural networks and the third model used networks with different topologies. The results of this study were impressive when two dermoscopy databases containing images from xanthous and Caucasian races were used for evaluation.

Ain et al. (2017) proposed a skin cancer detection system using Genetic Programming (GP) based feature selection from dermoscopy images. Their work extracted both high-level domain specific features recommended by the dermatologists and low-level LBP features. GP was used to identify the most significant features from the raw feature set to support subsequent benign and malignant cancer detection. The study by Esteva et al. (2017) performed skin cancer classification using deep neural networks. Their work employed an end-to-end convolutional neural network (CNN), where pixels and disease types were used as the training inputs. The CNN model was trained with a total of 129,450 clinical images with 2,032 different diseases. Evaluated against board-certified dermatologists on two binary classification tasks using clinical images, i.e., benign versus malignant (i.e. the deadliest cancer) and benign versus keratinocyte carcinomas (i.e. common cancers), their CNN model achieved a comparable performance with those of dermatologists.

Li and Shen (2018) proposed a deep CNN for the classification of melanoma. First, they segmented skin lesions using a fully convolutional residual network (FCRN). Then, by combining the FCRN and a multi-scale contextual information integration mechanism, they further enhanced the process. The segmented lesions' distinctive features were

extracted using a deep CNN model that consisted of more than 50 layers. To ensure that both networks did not overfit, they trained the networks through a residual learning mechanism. Their model was tested using the ISBI 2016 Challenge dataset which resulted in the model achieving the first place for classification and second for segmentation.

## 2.4 Feature Extraction

A sufficient number of features are needed to obtain an acceptable classification result. The dermatologist relies on skin lesion size, shape and colour variations for lesion diagnosis. Therefore, these features are vital for computerized diagnosis. Calculation of various shape features, e.g. asymmetry index (Stoecker et al., 1992, Lee, 1994), bulkiness score (Claridge et al., 1992), circularity factor (Cascinelli et al., 1992, Seidenari et al., 1995), fragmentation index (Green et al., 1994, Lee, 1994, Aitken et al., 1996), border structure and irregularity is required (Cascinelli et al., 1992, Seidenari et al., 1995). Additional eight standard ELM criteria have been employed to aid ELM-based diagnosis by Menzies et al. (2003), namely, black dots, radial streaming, brown globules, pseudopods, pigment network, overall depigmentation and pigmentation. These gradient values have been associated with a mean value and standard deviation in the work of Aitken et al. (1996) and Green et al. (1994), while fractal dimensions for the representation of border irregularity were determined by Hall et al. (1995). In general, a range of colour channels, i.e. standard deviation and the average value of the RGB (Lee, 1994, Green et al., 1994, Seidenari et al., 1995, Aitken et al., 1996), are used to constitute the colour descriptors.

Interpretation of lesion-related colours was undertaken by Cotton and Claridge (1996) based on an optical model of the skin. Their findings indicated that a two-dimensional surface patch in a three-dimensional colour space (CIE-LMS) contained all the normal skin colours. By contrast, the colours associated with abnormal skin structures were not included in this regular surface patch. Only a few studies have discussed parameters for defining dermoscopy structures and ELM criteria. Calculation of a feature for a lesion's reticular pattern (pigment network) was exemplified by Franke et al. (1993). The method of Grey-Level Co-Occurrence Matrix (GLCM) has been employed by Udrea and Lupu

(2014) for textural feature extraction. The study (Clausi and Zhao, 2002) has shown that a high G value beyond a given point made it more difficult to distinguish between disparity and contrast, even though the remaining measures were kept at an even level. Therefore, a uniform quantisation to 64 grey levels was performed to ensure that there were ample data within the matrix. The number of levels were chosen according to their experimental results.

## 2.5 Evolutionary Algorithms

The most popular heuristics algorithms are population-based optimisation models. These algorithms resolve diverse challenging optimisation problems efficiently. The procedures for resolving the optimisation problems include experimental settings (e.g. the pre-defined maximum number of iterations), multiple and random population-based search processes, and the fitness evaluation of the swarm. These algorithm processes are commonly divided into three categories, i.e. principles of physics, evolutionary and swarm intelligence-based algorithms.

The principles of physics have influenced variations of the algorithms, i.e. Gravitational Linear Search (GLS) (Webster and Bernhard, 2003), Gravitational Search Algorithm (GSA) (Rashedi et al., 2009), Charged System Search (CSS) (Kaveh and Talatahari, 2010), Artificial Physics Optimisation (APO) (Xie and Zeng, 2010) and Curved Space Optimisation (CSO) (Moghaddam et al., 2012). Evolutionary processes have also been the inspiration. The design and applications of evolutionary algorithms (EA) comprise several natural processes, such as mutation, selection, crossover, and reproduction (Back, 1996, Doerr et al., 2016), which are further enhanced in various ways to form popular models such as Evolution Strategy (ES) (Rechenberg, 1989), Genetic Programming (GP) (Koza, 1992), Genetic Algorithm (GA) (Holland, 1992), Differential Evolution (DE) (Storn and Price, 1997), Fast Evolutionary Programming (Yao et al., 1999), and Evolutionary Programming (EP) (Yao et al., 1999).

To address complicated issues and parallelism, (Holland, 1975, Holland, 1992) put forth the nature-inspired metaheuristic algorithm, GA, which is underpinned by the process of

natural selection propounded by the theory of biological evolution formulated by Charles Darwin. The GA has crossover, mutation and selection as its major operators. This algorithm is among the foremost evolutionary algorithms and is broadly applied, as highlighted in the literature. More specifically, issues associated with high dimensionality and feature selection are usually addressed via GA. Furthermore, different studies have developed numerous hybrid GA variant models so that the search space can be explored more efficiently. Besides GA, other swarm intelligent algorithms are especially introduced in detail below.

## 2.6 Swarm Intelligence Algorithms

The swarm intelligence algorithms are greatly inspired by the nature of socially collaborative animals and creatures. Such algorithms use a single individual in the swarm as a search agent and attempt to imitate the animals' flocking and schooling behaviours. Over the years, the following swarm intelligent algorithms have been proposed including, Cuckoo Search (CS) (Yang and Deb, 2009), bare bones PSO (Krohling and Mendel, 2009), Firefly Algorithm (FA) (Yang, 2010a), Bat Algorithm (BA) (Yang, 2010b), Particle Swarm Optimisation (PSO) (Kennedy, 2011), Ant Colony Optimisation (ACO) (Petmanson, 2011, Ünal et al., 2013), Directed Bee Colony Optimisation (DBC) (Kıran and Fındık, 2015), Dragonfly algorithm (DA) (Mirjalili, 2016), Moth-Flame Optimisation (MFO) (Mirjalili, 2015), Cultural Algorithm (CA) (Reynolds, 1999).

Artificial Bee Colony algorithm endeavours to reproduce the ingenious behaviour of the honeybee (Karaboga and Basturk, 2007). It is designed to address optimisation issues by mimicking the natural foraging behaviour of honeybees, with food sources and nectar respectively denoting solutions and fitness values. The bees are divided into three categories: the scouts, onlookers and employed. The employed bees know of food sources and can determine where nearby ones will be, and, upon return to the hive, they begin an intricate and precise dance. It is the onlooker's job to monitor this performance, as the precise movements of the bee indicate the exact locations of accessible food sources, and the quantity of food available. When a food source is abandoned, the employed bee then ventures forth in search of a new one and becomes a scout. The original ABC algorithm

has been amended numerous times over the years to overcome limitations of the original ABC algorithm. Similarly, the Colonies of ants, as imitated by Ant Colony Optimisation (ACO) (Petmanson, 2011, Ünal et al., 2013), can determine the ideal route from their prime source of food to their nests using ascent or pheromone trail.

The behaviour of a typical grey wolf pack is the inspiration for the Grey Wolf Optimizer (GWO) algorithm (Mirjalili et al., 2014b). Grey wolves have a system of hierarchy within their groups, which is especially prevalent when hunting prey. The algorithm replicates three main stages of the hunt: locating, encircling, and then attacking. The algorithm has been employed by numerous studies for solving diverse optimization problems.

One of the most popular swarm-based algorithms is the PSO algorithm, owing to its simple search concepts and its low cost to implement. Most of the other algorithms are directly or indirectly developed based upon the primary search processes of PSO. The PSO engages in the simulation of the way in which organisms socially interact, as exemplified by the collective movements of migrating birds or schooling of fish (Kennedy, 2011). Within PSO, every possibility is examined as a separate particle within a certain search space, and each particle possesses a distinct position and velocity. As motion takes place, the position of each particle changes because of its velocity and the way in which its counterparts are behaving, and this process continues until the swarm establishes an optimum position. Although PSO has its drawbacks, such as premature convergence, other strategies can further enhance or amend its key processes to counteract the issue.

There are many different ways to solve the premature convergence challenges of conventional PSO (Zhan et al., 2011, Xue et al., 2013), and some of these were discussed in detail in existing literature. Especially, in recent years, PSO variant models have been widely employed for solving diverse optimization problems. These works have placed increasing emphasis on PSO and its modified models because of their clarity, population-based and fast convergence nature. Recently, numerous PSO based methods have proposed for solving the feature selection issues, such as chaotic binary PSO (Chuang et al., 2011a), Taguchi chaotic binary PSO (Chuang et al., 2011b), discrete PSO (Unler and

Murat, 2010) and geometric PSO (Talbi et al., 2008). Other modified models of the original PSO algorithm include, e.g. MS-PSO (Yang et al., 2017), BBPSOV (Srisukkham et al., 2017) and GPSO (Chen et al., 2016). We discuss the PSO-based feature selection models in detail below.

Nasir et al. (2012) proposed a PSO variant called the Dynamic Neighbourhood Learning Particle Swarm Optimizer (DNLPSO). It employed a neighbourhood-based learning strategy where the neighbourhood historical best information was used for velocity updating of the current particle. The neighbourhoods were also updated and reconstructed dynamically after a certain number of iterations to preserve swarm diversity.

An advanced PSO framework, known as Enhanced Leader PSO (ELPSO) (Jordehi, 2015), has been demonstrated to mitigate premature convergence of the original PSO model. ELPSO extended the diversity of the swarm by employing the following mutation strategies for swarm leader enhancement, i.e. Cauchy, Gaussian, opposition centric, and differential evolution (DE) mechanisms.

Multi-runs of standard binary PSO were undertaken in the work of Krisshna et al. (2014). The resulting model was known as ThBPSO. The best solution, i.e. *gbest,* detected from each run was retained. A threshold subsequently enabled to establish how important every dimension of the universal optimal solutions was. If the overall number of selections of a feature in the previous runs exceeded the pre-established threshold, then that feature was chosen and attributed as importance. Seven benchmark datasets were employed to test the system, which was proven to perform better compared to other advanced approaches.

Zhang et al. (2015a) developed a binary Bare Bones Particle Swarm Optimisation (BPSO) model for feature selection. To cultivate the optimum local particles, their binary BPSO was dependent upon a robust system of memory, which picked and prioritised personal best particles to provide the swarm with an increase and enhancement in variation. Overall, their model combined a reinforced memory strategy and a uniform combination to increase local and global search capabilities. Wang et al. (2013) proposed an algorithm named as Gaussian bare-bones differential evolution (GBDE). It employed Gaussian

distribution as its primary mutation method and relied on a self-adaptable approach to maintain crossover prospect modifications. The combined model produced a significantly faster rate of convergence. The improved approach provided superior results, even when compared with bare bones and DE variant algorithms.

Additionally, Lim and Isa (2015) developed Adaptive Division of Labour PSO (ADOLPSO). In their work, new particles were produced based on the memory of the swarm using the reflectance and convex operators. With the use of the Comprehensive Learning PSO (CLPSO) (Hu et al., 2014), particles were invited to learn from the best solutions of other neighbouring particles in each dimension.

Mahmoodabadi et al. (2014) proposed a PSO variant, referred as high exploration PSO (HEPSO). HEPSO combined PSO with a multi-dimensional GA function and the food source identifier of bee colony optimisation. This merge enhanced the search and population diversity. Ultimately, when analysed and assessed, according to benchmark performance indicators, HEPSO was seen to be the most efficient among the selected the PSO variants.

Nevertheless, Chuang et al. (2011a) proposed a chaotic binary PSO (CBPSO) for feature selection. It involved the integration of chaotic maps (Tent and Logistic maps) with binary PSO, to avoid local optima traps. According to observed outcomes, merging Tent maps with the binary PSO achieved the best performance.

To create a Pareto front of non-dominated resolutions, Xue et al. (2013) suggested two PSO-based multidimensional feature selection algorithms, i.e. crowding, mutation, non-dominated sorting PSO (NSPSO) and dominance PSO (CMDPSO). The former brings together PSO with methods of mutation, dominance, and crowding, as implied. The latter combines principles of PSO with non-dominated sorting. The two algorithms create a crowding distance for non-dominated resolutions and help to preserve the variance and quality of the finest particles.

The work of Li et al. (2015) suggests another hybrid PSO algorithm. It relies upon a weighted particle and fuzzy logic to direct the swarm. The weighted particle determines the direction, but different criteria such as inertia weight and attraction levels are responsible for regulating local exploitation and global exploration. It has been extensively tested with the use of ten benchmark indicators and a non-linear neural network system.

Lynn and Suganthan (2017) proposed an ensemble of PSO algorithms. Their ensemble model utilized a pool of PSO variants including inertia weight PSO, comprehensive learning PSO (CLPSO), and distance-based locally informed PSO (LIPS). The method divided the swarm into two subpopulations, i.e. one large and one small subswarms. CLPSO was applied to the small subpopulation to maintain swarm diversity whereas a self-adaptive probabilistic selection scheme was used to identify the best PSO algorithm from the pool in each iteration using the large subswarm. In their method, if a PSO variant was able to achieve performance improvements within a fixed learning period, it was stored in a success memory, otherwise it was recorded in the failure memory. The PSO strategies with higher success rates were more likely to be selected to guide the search of the large subswarm.

Gou et al. (2017) developed a PSO variant based on individual difference evolution. Motivated by the social and psychological models, their work treated each particle as a virtual human, and proposed an emotional status fitness indicator for each particle. It not only determined a specific evolutionary mechanism for each particle based on its emotional status and current fitness, but also utilized a re-starting mechanism to increase swarm diversity. Optimal parameter settings were also explored in their work.

Dash et al. (2017) developed two hybrid evolutionary models, known as CSPSO and ICSPSO, respectively. The search strategies of PSO and Cuckoo Search (CS) were integrated to increase local exploitation and global exploration of the original PSO algorithm. Both proposed methods enabled each particle to conduct position updating using the original PSO mechanism. A new nest was then generated using the Levy flight operator in CS from the position vector of a current particle if this particle was ranked

within the best 4% of the population. In ICSPSO, the mutation and crossover operations of Differential Evolution (DE) were integrated with CSPSO to further enhance the global best solution in each iteration. Both algorithms were used for the design of linear phase multi-band stop filters and the search of the desired impulse responses of the filters.

Ouyang et al. (2017) proposed another PSO variant, known as an improved global-best-guided PSO with learning operation (IGPSO), for solving engineering design optimization problems. In their work, the overall swarm was divided into three categories, i.e. current population, historical best population, and global best population. A specific search strategy was dedicated to each subpopulation. Besides using the original PSO operation to guide the search of the current population, a Gaussian distribution based local exploration mechanism was applied to the historical best subswarm to enable each personal best particle to learn from other more promising historical best particles independently. Stochastic learning and opposition based learning operations were also used to further enhance the global best solution and to overcome the local optima traps. IGPSO outperformed other PSO variants and classical search methods when evaluated using a total of 25 unimodal and multimodal benchmark optimization functions.

On the other hand, an autonomous group PSO (AGPSO) model was developed by Mirjalili et al. (2014a). Adaptively decreasing cognitive parameter and increasing social parameter were employed to balance between exploitation and exploration. Chen et al. (2017) proposed a biogeography-based learning PSO (BLPSO) model to further enhance CLPSO. Both CLPSO and BLPSO performed velocity updating of each particle using the personal best solutions of different exemplar particles for different dimensions. In comparison with CLPSO where learning probability and tournament selection were used to generate exemplars for each particle, BLPSO employed the migration of biogeography-based optimization for exemplar generation. BLPSO achieved impressive performance for the evaluation of diverse challenging benchmark functions.

There are other algorithms which attempt numerous ways to enhance the PSO model (Zhan et al., 2011, Chang, 2015, Lynn and Suganthan, 2015, Zhang et al., 2015a, Srisukkham et al., 2017, Mistry et al., 2017, Yogesh et al., 2017). Modifications have

been made to the quadratic interpolation processes (Beheshti and Shamsuddin, 2015), the positional updating equations (Yang et al., 2017), centripetal acceleration (Beheshti and Shamsuddin, 2014), and various strategies (Qin et al., 2015, Tang et al., 2015) and numerous subpopulations (Chang, 2015). Further alternatives present approaches to resolve the discrete optimisation problems, yet more try to produce stronger, more powerful versions by focusing on the inertia weight update strategies (Shi and Eberhart, 1998, Chen et al., 2013, Taherkhani and Safabakhsh, 2016), while others attempt to incorporate more adaptability.

## 2.7 Classification

The reduced or selected features are used as the input for recognition and classification methods. In this procedure, features from the feature extraction or feature selection process are collected. They are prepared in two groups, which are training and test data sets. The normalisation process is used for regulating datasets suitable for a supervised neural network structure (Jain et al., 1996). The classification process then uses neural network or other machine learning methods to recognise information/pattern embedded the data and classify the input data into class types for analysis and decision-making. In this section, some advanced machine learning approaches are discussed.

### 2.7.1 Support Vector Machines (SVM)

SVM contains a very sophisticated learning process that still has a simple mentation. SVM is a supervised learning method that is used to solve regression (Support Vector Regression) and classification (SVM) problems. Most of the classifier-based knowledge transfer work has been constructed using the original SVM classifier. SVM is performed by creating the *n*-dimensional hyperplane that separates data into two classes. The goal of SVM is to locate an optimum hyperplane that could separate clusters of vectors, which puts one category of the target variable on one side and the other class on the other side. The support vectors are the vectors near the hyperplane. By using polynomial functions, the best classification rates have been achieved in many existing works.

Classification and regression issues are typically addressed via the supervised learning models such as SVM or neural networks (Hearst et al., 1998). Based on a series of training examples with labels, which indicate the category that the examples belong to, a model adopting the linear or non-linear function is developed by the SVM training algorithm to allocate new examples to specific classes. Therefore, SVMs are suitable for conducting not only linear classification, but also non-linear classification through implicit mapping of their inputs into high-dimensional feature spaces based on the kernel trick.

### 2.7.2 Ensemble Classifiers

As explained by Rokach (2010), models consisting of several weaker and separately trained learners (i.e. base models) are known as ensemble methods. In this context, the final prediction is the outcome of the combination of weak learners' predictions. Two objectives must be accomplished to develop an ensemble classifier, namely, diversification of weak learners and creation of a more effective mechanism to enhance final precision by merging weak learners' decisions. Popular ensemble methods include AdaBoost, Stacked Generalization, Boosting, Gradient Boosting Machines and Random Forest.

### 2.7.3 Convolution Neural Networks

The Convolutional Neural Network (CNN) is a popular approach for image processing. It is part of back propagation neural network that will create a connection between standard adaptive filters and feed-forward neural network to operate on images (Sahiner et al., 1996). In comparison with traditional neural networks, the CNN model with possible non-linear activation functions with or without down sampling performs feature learning (Browne et al., 2008). A fully connected layer is used to perform the classification with machine learned features as inputs. The CNN model has been used for diverse medical imaging, large-scale object recognition and scene classification applications.

## 2.8 Conclusion

Skin cancer detection is a multi-stage process to identify benign and malignant skin lesions, including pre-processing, for example, using dull razors and median filters to remove hair and other noise. Then, the images were segmented using a pixel limitation technique to separate lesions from the skin background. Feature extraction is subsequently conducted. The features extracted by our system reflect the well-known asymmetry, border irregularity, colour variegation and diameter (ABCD) of dermatology criteria. After extracting thousands of raw shape, colour and texture features from the lesion areas, feature optimization methods are used to identify the most discriminating significant feature subsets for healthy and cancerous lesions. And finally, multiple classification methods, such as Support Vector Machine and ensemble models, have been employed to perform benign and malignant lesion recognition.

The overall process of the identification of benign and malignant skin lesions from dermoscopic images, especially the feature selection process, is motivated by various related research in the field. As an example, in our work, multiple optimization algorithms are proposed which offer the prospect of achieving reduced computation cost and at the same time enhanced accuracy rates for the classification of lesion images. Overall, there are three most closely related works that motivated this research for evolutionary algorithm based feature selection. Proposed by Krohling and Mendel (2009), the BBPSO model employed chaotic accelerated attraction and enemy avoidance operations with the consideration of the local and global best and worst experiences to increase search diversity. But this model has limited discriminative capabilities owing to excessive following of the best leaders, it tends to select a high number of features and shows limited capability when solving complex optimization problems, evidenced by our experiments.

Proposed by Krisshna et al. (2014), the ThBPSO model used multiple runs of the original PSO model to generate and identify the final best solutions. Because of using the original PSO model mainly, the ThBPSO method has a higher probability of premature convergence and being trapped in local optima. Finally, the ELSPO model (Jordehi, 2015) included several local exploitation strategies for swarm leader enhancement. The drawback of ELPSO is that it only enhanced the global best solution without increasing

the diversity of the overall swarm. This research presented in this thesis has been motivated by the above three and other developments in the field by embedding and proposing multiple search mechanisms to increase search diversity and efficiency, to overcome premature convergence, of the original PSO algorithm.

# CHAPTER 3
# PROPOSED METHODOLOGY 1 WITH PSO AND GA-BASED FEATURE SELECTION

It is challenging to develop an intelligent agent-based or auto-detection system to conduct long-term automatic health monitoring. To develop a disease diagnosis system in the real world, certain requirements such as robustness and efficiency must be given priority. This research aims to deal with such challenges by presenting an intelligent decision support system for skin lesion recognition as the initial step. The proposed method could also be embedded in an intelligent automation service for health monitoring in domestic environments to promote early diagnosis. The system is developed to identify benign and malignant skin lesions using process with multiple steps, including pre-processing for noise and hair removal, segmentation, feature extraction, feature selection and classification. After extracting thousands of raw features relating to shape, colour and texture from lesion areas, two separate optimisation methods, the GA and PSO are used to identify the most discriminating feature subsets for healthy and cancerous cases. The classical PSO and GA models are selected owing to the fact that they carry out distinctive search strategies to each other. They are selected also because of their capabilities and efficiency in solving diverse optimization problems and simplicity for implementation. Specifically, PSO follows the swarm leader to explore the search space, which has better global exploration capability, while GA is based on mutation and crossover to generate offspring solutions, therefore having a better capability for local exploitation. At this initial stage, the research mainly focusses on these classical models for initial experiments. Moreover, owing to the fact that there are various limitations in these conventional methods, such as local optima traps and lack of search diversity, several

improved hybrid models are proposed and presented in subsequent chapters. After performing feature selection using PSO and GA, an SVM classifier has been employed to recognize benign and malignant lesions. The empirical results indicate that the proposed approach achieves superior performance compared to other methods reported in the literature. The comparison is performed by evaluating 1,300 images from the Dermofit dermoscopy image database. The features employed in the present study have high-dimensionality characteristics. However, not all features of skin lesion have the same level of importance. The selection of features has implications for the determination of the nature of a lesion as benign or malignant, but this may be inaccurate if unnecessary features are used. Future regression or classification models could benefit from a reduced series of essential features that could make them more precise and comprehensive. Texture and colour are complex properties, and it is not well understood how important different levels of sharpness and colour are, even though differentiating features could be distinguished based on domain knowledge. Therefore, feature selection must be performed so that only the most important features are selected.

This research proposes an intelligent decision support system for the classification of benign and malignant skin lesions using GA and PSO-based feature selection. As shown in Figure 3-1, there are five principal processes in the proposed system for the identification of skin cancer; namely, pre-processing, skin lesion segmentation, and feature extraction, and then most importantly, the PSO and GA are utilised to perform feature selection and so as to recognise the most discriminative features for the classification of benign and malignant skin lesions. These evolutionary algorithm based feature selection methods can eliminate unnecessary features and thus enhance the accuracy of classification. Subsets of features, i.e. the most distinctive features, will eventually be employed as input for the SVM to enhance classification results in dynamic environments. The proposed methodologies on the first feature selection model are discussed in detail below.

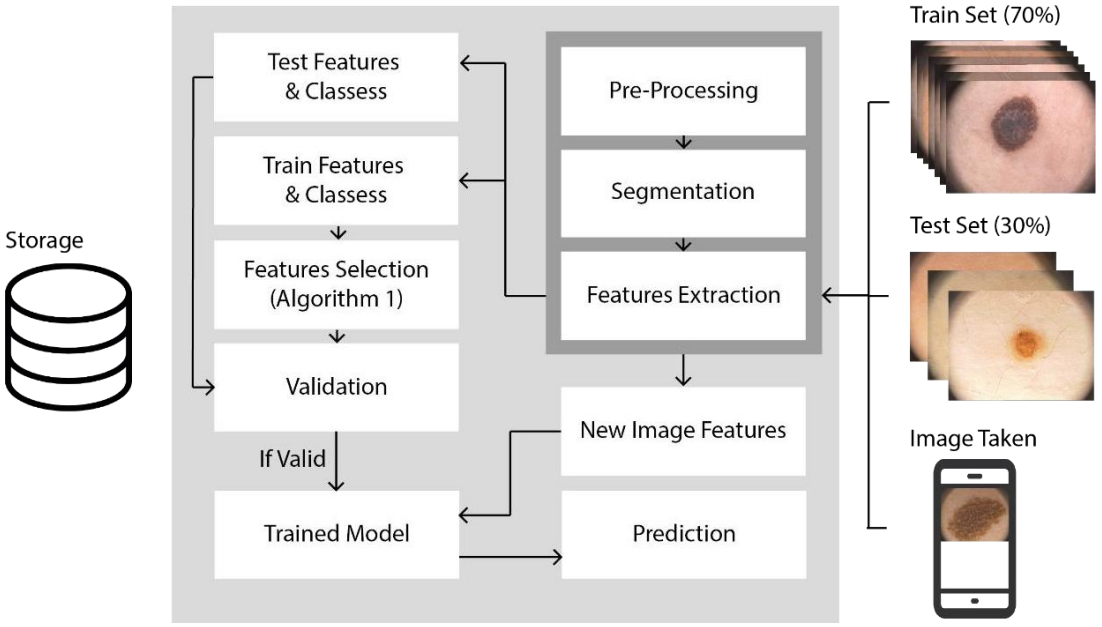



**Figure 3-1: Machine learning procedures for skin lesion classification**

## 3.1 Pre-processing

A pre-processing method is essential before feature extraction when classifying benign and malignant lesions and this involves transforming raw data into an understandable format for further processing. In real-world applications, data are often unreliable, inadequate and may contain noise that affects detection performance. Removing any noise and irrelevant marks which confuse the classifiers is required. The following pre-processing steps are conducted in this research. Noise filtering, image segmentation, and grayscale conversion are initially conducted before subjecting image of lesions to classification procedures. For instance, a hair removal technique is undertaken to reduce the effects of thin hairs and a median filter is used to eliminate ‘salt and pepper’ noise. Subsequently, the lesion is separated from the background by removing the skin through a ground truth motivated image segmentation algorithm. Finally, the original RGB images are then converted into grayscale images.

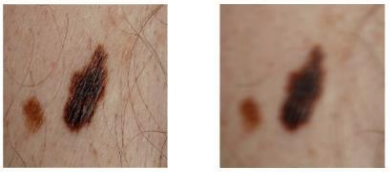


**Figure 3-2: Before (Left) and after (Right) hair removal**

The Enhanced Dull Razor algorithm (Kiani and Sharafat, 2011) was used to remove hairs from images where morphological closing image processing was generalised to grey-level images, followed by the identification of the thin hair with an extended hair outline. Bilinear interpolation was implemented to substitute the identified pixels of the hairs. The MATLAB 'roifill' function was also used to perform a filling operation using an interpolation method based on Laplace's equation. This step resulted in a smooth fill inward from the borders of the region of interest, as shown in Figure 3-2. Subsequently, the clarity of the image was enhanced by improving its shape and edges. Image borders were sharpened using contrast enhancement. Moreover, this process may also improve the quality of the segmentation.

## 3.2 Segmentation

Image segmentation is a technique used to determine the shape and size of the border. Segmentation can also separate the object from its background based on the features needed for extraction (Neoh et al., 2015). After removing noise and hairs from the image, the lesion needs to be separated from the skin, and therefore the analysis leading to diagnosis is conducted using only the critical area. Previous studies have proposed several different types of segmentation methods which give high accuracy, such as region-threshold and edge-based methods (Zhang et al., 2013). The Adaptive Snake (AS) approach has been chosen for this research, because of its efficiency as indicated in previous research. Silveira et al. (2009) chose several existing segmentation methods for comparison, including the adaptive threshold (AT), adaptive snake (AS), gradient vector flow (GVF), Chan's level set method (C-LS) and fuzzy-based split and merge (FBSM). Each method has been evaluated for segmentation performance by comparison of the results with the ground truth established by the dataset provider. Based on the results, AS

is efficient for establishing a discriminatory analysis that divides the image into two classes of pixels. In the first instance, the chosen colour image is rendered in monochrome. Then corresponding threshold limits are set within the grey spectrum, and the pixels that occur within the range set by the limits are selected. Following this, non-lesion pixels are assigned a value of zero. To extract multiple features such as colour and area, the segmentation results of this threshold-based method are plotted on multiple images. Figure 3-3 shows a segmented skin lesion using one of the images extracted from the dataset used in the present studies.

**Figure 3-3: Segmentation results: original lesion image (left), segmented cropped image (middle), segmented binary image (right)**

## 3.3 Feature Extraction

The identification of regions consisting the internal and external parts of lesion enables the creation of an asymmetry index, since asymmetry is highly important in distinguishing between malignant and non-malignant lesions. Meanwhile, slight edge irregularities are detected based on a border irregularity index. In addition to these morphological aspects, colour features such as relative chromaticity, ratios of red, green and blue, and factors associated with the tone of the lesion are extracted based on the HSV, RGB and Lab colour spaces. Meanwhile, after segmentation, image features are extracted for subsequent classification.

**Table 3.1 : Features extracted from a skin lesion image**

| Features# | Feature List | Feature Types |
|---|---|---|
| f1 | Border irregularity index | Shape |
| f2 | Perimeter | |
| f3 | Solidity | |
| f4 | Extent | |
| f5 | Equivalent diameter | |
| f6 | Asymmetry index | |
| f7 | Form factor | |
| f8 | Roundness | |

| | | |
|---|---|---|
| f9 | Major and minor different | |
| f10 | Compactness index 1 | |
| f11 | Compactness index 2 | |
| f12 | Relative chromaticity | Colour (RGB, HSV, Lab) |
| f13 | Lightness difference | |
| f14 - f16 | Colour difference | |
| f17 | Colour ratio | |
| f18 - f20 | Colour mean | |
| f21 - f23 | Colour standard | |
| f24 - f26 | | |
| f27 - f3914 | Generalised Co-Occurrence Matrix | 12 GCM features, 3 grey quantisation level x 6 inter pixel distances in 3 colour spaces (RGB, HSV, Lab) & 6 colour pairs (RR, RG, RB, GG, GB, BB) |

Several methods have been identified for feature extraction. Overall, most of the relevant work has employed the ABCD rules of dermatology for feature extraction. In this research, several types of measurements such as compactness index, fractal index, and edge abruptness are used to indicate border irregularities to identify these features which will be included in this application. In the skin automated diagnosis system, feature extraction is constructed based on ABCD rules by dermatology. The three measures of shape, colour and texture provide the following features as shown in Table 3.1.

### 3.3.1 Shape Features

**Asymmetry:** A melanocyte lesion may be diagnosed according to several identifiers, of which one of the most significant is a lack of symmetrical morphology. In dermatological terms, the ABCD rule model rates this aspect as the most crucial factor. In consideration of the symmetry feature, many factors are relevant concurrently, including colour, texture and morphology. A three-fold classification system can be derived from the measurement of symmetry, with three classes of outputs representing total symmetry, lack of symmetry along a single axis and lack of symmetry along dual axes. Lesion asymmetry is evaluated by calculating the area of inner and outer of the lesion, using the following formula:

$$AI = \frac{\Delta AK}{AL} * 100 \qquad (3.1)$$

where $AI$ represents the asymmetry index. $\Delta AK$ represents the area between the two halves of the lesion, and $AL$ denotes the lesion area.

**Border irregularity:** Irregularities occurring along the edge of a malignant lesion offer useful information concerning that lesion's nature. Typically, the edge of a malignant lesion usually exhibits four factors of interest which are density, fractal dimensions, radial variability and the extent to which its contour exhibits small irregularities. To identify the irregularity of a lesion border, each lesion is divided into four areas for comparison: left, right, below and above. From this, the following four border irregularity indexes are as shown in Figure 3-4.

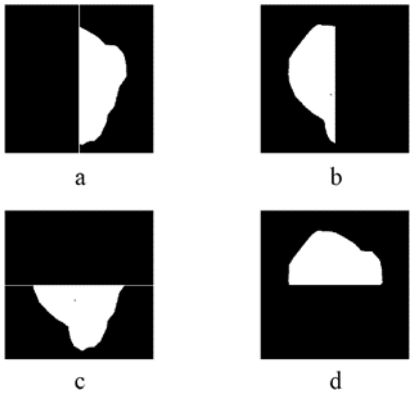



**Figure 3-4 : Calculating the border irregularity index, (a) border right, (b) border left, (c) border-bottom, (d) border upper**

$$I = \frac{ab}{2\pi(a^2 + b^2)} \frac{P^2}{\Delta A} \qquad (3.2)$$

where $I$ represents an irregularity with $a$ and $b$ representing the lengths of major and minor axes of the lesions, $P$ represents the perimeter of the lesion and $\Delta A$ indicates the area of corresponding.

**Compactness**: Another relevant feature is the degree to which the lesion can be described as compact. To determine this aspect, a comparative analysis is performed between the lesion's boundary and a circle with a circumference of the same length. It is the former of these two numerical values that presents a challenge in its assessment. One solution to this issue is to use the proportions of the most easily measured values of maximum and equivalent lesion diameters as defined in Equation (3.3), where $C$ represents compactness.

$$C = \frac{4\pi P^2}{\Delta A} \tag{3.3}$$

### 3.3.2 Colour Features

The range of colour types utilised in diagnosing a melanocyte lesion can be broadly categorised into black, grey-blue, brown (dark), brown (light), red and white, which are indicators of a malignant skin lesion when all are present. Dermatological analysis allows for the determination of whether a colour category exists in an image and, if so, where it exists. This positional information is noted via a binary mask application, with image segmentation performed by the dermatological professional (see Figure 3.5 as an example with separate colour categories being present). In this chapter, three types of colour space are used: HSV, RGB and LAB.

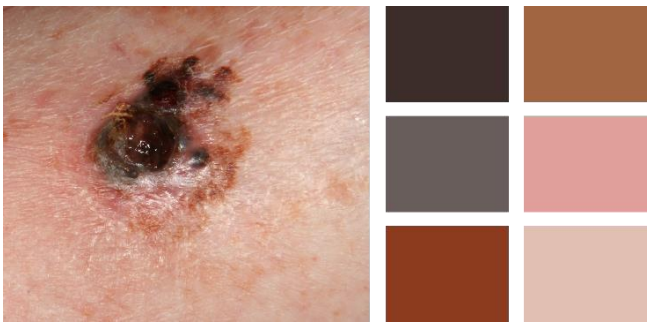


**Figure 3.5 : Colour variations of a malignant skin lesion**

The relative chromaticity of red, green, and blue in a skin lesion is also significant. The chromaticity of red, green and blue is expressed as follows:

$$Rel_{Chroma,Red} = \frac{r_L}{r_L + g_L + b_L} - \frac{r_S}{r_S + g_S + b_S} \tag{3.4}$$

In the above expression, $r_L$, $g_L$ and $b_L$ represent the mean red, green and blue appearing within the lesion respectively. The chromaticity of green and blue in relation to red/blue and red/green respectively are expressed in a similar way. In the case of red, the ratio represents the average of the red constituent present in a lesion divided by the mean colour of the surrounding non-lesion skin. The ratio of red is expressed as follows:

$$Ratio_{Red} = \frac{r_L}{r_S} \qquad (3.5)$$

The ratios for green and blue are defined similarly. Finally, the factors exhibited regarding the lesion's tone, in term of chroma, colour and hue are examined in relation to the non-lesion skin. The lesion's tone, chroma and hue are represented by $L^*$ , $C^*$ and $H^*$ respectively in the following equations:

$$(\Delta L^*) = L^*_{Lesion} - L^*_{Skin}$$
$$(\Delta C^*) = C^*_{Lesion} - C^*_{Skin}$$
$$(\Delta Clr^*) = \sqrt{(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2}$$
$$(\Delta H^*) = \sqrt{(\Delta Clr^*)^2 + (\Delta L^*)^2 + (\Delta C^*)^2}$$

where

$$\Delta a^* = a^*_{Lesion} - a^*_{Skin}$$
$$\Delta b^* = b^*_{Lesion} - b^*_{Skin} \qquad (3.6)$$

where $a$ and $b$ are the colour-opponent dimensions of the Lab colour space.

### 3.3.3 Texture Features

The texture of a lesion can be estimated by many objective measures derived from generalized co-occurrence matrix (GCM). Through a body of existing research Ballerini et al. (2012), the grey-level co-occurrence matrix (GLCM) has been extensively used and is a widely-adopted and popular methodology. GLCM provides several measures for statistical assessment which are employed in this research, and each is grey-level shift-invariant in nature. These measures enable the sensitive recognition of linear shift regarding the intensity of illumination, such that texture can be categorised in these terms. Previous research (Clausi and Zhao, 2002) has demonstrated that a point exists beyond which a high G value leads to a reduced ability to differentiate in terms of disparity and contrast, despite maintaining even levels of the other measures. To populate a matrix with a sufficient amount of data, an equal quantisation to 64 grey levels was carried out, with this number being above a lower bound of 24 which was selected based on the findings

of existing research (Ballerini et al., 2012). Such low values also minimise the impact of noise in the image.

It is recommended that the GLCM after normalisation presents a stable level of density, to provide confidence in statistical estimation within the joint probability distribution. In this research, the measures are taken in three colour spaces, i.e. RGB, HSV and Lab to reduce the impact of differences in lighting before the colour extraction, plus six colour pair, i.e. RR, RG, RB, GG, GB and BB. Three grey-level quantisation at 64, 128, 256, are used for every colour space. The twelve texture features represent autocorrelation, correlation, cluster prominence, dissimilarity, entropy, energy, maximum probability, contrast, homogeneity, cluster shade, inverse difference moment and variance as mentioned by Haralick (1979), and six inter-pixel distances. Eventually, each image yields 11 morphology-based features and 3888 textural features, giving a total of 3914 features.

## 3.4 The Proposed PSO and GA Search Strategies for Feature Selection

The features listed in Table 3.1 are high dimensional and are not all equally important for the identification of benign and malignant lesions. Therefore, feature selection is conducted to identify the most significant features. In this research, PSO and the GA are employed for feature selection because of their superior search capabilities (Zhang et al., 2015c, Mistry et al., 2017). The GA employs crossover, mutation and replacement functions to generate better offspring and thus the most discriminative features can be identified. In this work, both algorithms reduced the number of feature dimensions, from 3914 to 1472 for the GA, and from 3914 to 1254 for PSO. Subsequently, the most discriminative feature subsets are identified and used as inputs to the SVM classifier for the classification of healthy and cancerous skin conditions. The pseudo-codes of the PSO and GA are also provided below.

**Algorithm 3.1: Genetic Algorithm**

| | |
|---|---|
| **1** | **Start** |
| **2** | Initialize a population randomly (e.g. 50 particles); |
| **3** | Evaluate the population to identify the initial best leader**,** $gBest$; |
| **4** | **While** (! Stop condition) do {// 500 iterations |
| **5** | (P^' (t)) ← Variation (P(t)); //Creation of new solutions |
| **6** | Evaluate population (P^' (t)); //Evaluate the new solutions |
| **7** | P(t+1) ← Apply genetic operators (P^' (t)∪Q); |
| **8** | t ←t+1; |
| **9** | } **End While** |
| **10** | **End** |

**Algorithm 3.2: Particle Swarm Optimisation**

| | |
|---|---|
| **1** | **Start** |
| **2** | Initialize a population randomly (e.g. 50 particles); |
| **3** | Evaluate the population to identify the initial best leader**,** $gBest$; |
| **4** | **While** (! Stop condition) do {// 500 iterations |
| **5** | **For** (each particle) do { |
| **6** | **If** (f(xi) < $pBest$) do { |
| **7** | $pBest$ = xi; |
| **8** | } **End If** |
| **9** | **If** ($pBest$ < $gBest$) do { |
| **10** | $gBest = pBest$; |
| **11** | } **End If** |
| **12** | } **End For** |
| **13** | **For** (each particle) do { |
| **14** | **For** (each dimension) do |
| **15** | Search using Eq. (3.7) & (3.8) in each dimension; |
| **16** | } **End For** |
| **17** | } **End For** |
| **18** | } **End While** |
| **19** | **End** |

A swarm-based algorithm PSO is often employed for feature selection. PSO shows outstanding search abilities in addressing a wide range of issues of optimisation and it performs the search process using personal and global best solutions. The PSO operations for velocity and position updating are defined by Equations (3.7) and (3.8).

$$v_{id}^{t+1} = w \times v_{id}^{t} + c_1 r_1 (p_{id} - x_{id}^{t}) + c_2 r_2 (p_{gd} - x_{id}^{t}) \tag{3.7}$$

$$x_{id}^{t+1} = x_{id}^{t} + v_{id}^{t+1} \tag{3.8}$$

where the positions of particle $i$ in the $d^{th}$ dimension in repetitions are respectively indicated by $x_{id}^{t+1}$ and $x_{id}^{t}$, while the velocities in the $t$+1th and $t^{th}$ repetitions are respectively indicated by $v_{id}^{t+1}$ and $v_{id}^{t}$. The inertia weight for embedding the repetition effect of the preceding velocity is indicated by $w$, whereas the acceleration coefficients are represented by $c_1$ and $c_2$, with $r_1$ and $r_2$ constituting random vectors. The personal best solution of particle $i$ is denoted by $p_{id}$, while the global best solution in the $d^{th}$ dimension is denoted by $p_{gd}$. In this research, the features selected are represented in a set of $n$ binary digits where serves to represent a feature subset. The values 0 and 1 correspond as removed and selected features. For instance, a particle with the binary set of 01010010 suggests that the second, fourth and seventh features were chosen. A set of candidate solutions depicted as a particle within a swarm in solving an optimisation issue is evolved iteratively to achieve enhanced resolutions, as shown in Equation (3.9). The overall process terminates when the maximum number of iterations is reached, or a satisfactory outcome has been achieved. The fitness function of each particle is composed with two aspects based on the classification accuracy and the number of selected features, as defined by the following equation:

$$fitness(C) = w_a * Accuracy + w_f * (numberFeatures)^{-1} \quad (3.9)$$

where $w_a$ and $w_f$ represent the weights for classification accuracy and the selected features respectively. A higher value is assigned to $w_a$ (e.g. 0.9) as opposed to $w_f$ (e.g. 0.1), owing to the higher priority for the classification performance as compared with the number of selected features.

## 3.5 Classification

Selected features are used for the recognition and classification of benign and malignant lesions. A wide range of classifiers is explored. Eventually, the SVM classifier was used for the classification of benign and malignant skin lesions because of its superior performance (Gilmore et al., 2010). The SVM is a supervised learning method used to solve regression and classification problems (Zhang et al., 2013, 2015b). The SVM

creates an *n*-dimensional hyperplane that separates the data into two classes. The goal of the SVM is to pinpoint an optimum hyperplane which can separate clusters from vectors to categorise the target variables on one side and the other class on the other side. The support vectors refer to the vectors near the hyperplane. Moreover, the SVM is regarded as an inherently straightforward application, with fewer parameters needed for tuning, a high capacity to generalise output within pattern-recognition scenarios and a strong potential to accommodate input information of higher dimensions. In this research, an SVM with radial basis function (RBF) as well as optimal parameter settings is utilised for classification.

## 3.6 Evaluation

This research employs the Dermofit dataset for the evaluation of the proposed system. This dataset consists of images of 1,300 lesions including 850 benign and 450 cancer images. Most of the 1300 images are 8-bit red-green-blue (RGB) colour images in a variety of resolutions. Since most of them, however, consist of 720 pixels in the longest dimension, other images that do not meet this criterion were resized to 720 pixels in the longest dimension with the aspect ratio being maintained. In each case, the lesion appears in the centre of the image with non-lesion skin visible in the corners of the image.

Twelve hundred images were employed for training and one hundred images for testing from the above database. Although PSO and the GA are comparatively superior in the outcomes they deliver, the experiments also revealed limitations due to the difficulty of finding global maxima which will be the subject of further research. In this research, the PSO and GA approaches have been applied in two separate experiments to select relevant and reliable factors to carry out skin lesion categorisation. The results of the experiments were compared with those obtained from related research, such as the study by Ayoub et al. Ayoub et al. (2012). The following formulae were used to calculate accuracy, sensitivity and specificity in the evaluation of the system.

$$Accuracy = \frac{(Total\ of\ detected\ lesions)}{(Total\ dataset\ images)} * 100 \tag{3.10}$$

$$Sensivity = \frac{TP}{TP + FN} \quad (3.11)$$

$$Specificity = \frac{TN}{TN + FP} \quad (3.12)$$

In these formulae, TP is the number of true positives, with FN as the number of false negatives, TN the number of true negatives, and FP the number of false positives. Sensitivity and specificity respectively quantify the percentage of classifications relevant to benign and malignant lesions. The experiments use 100 images depicting 50 malignant and 50 benign lesions. The results obtained from the full set of images are set out in the table below. The results show that model obtains promising performance, with average accuracy rates of 92.11% and 90.93% for GA and PSO respectively.

**Table 3.2: Evaluation Results**

| | **Methodology** | **Sample size** | **Accuracy (%)** |
|---|---|---|---|
| Gilmore et al. (2010) | SVM | 199 | N/A |
| Barhoumi and Baâzaoui (2014) | ROC | 122 | 85.00 |
| Alcón et al. (2009) | CFS + LMT | 152 | 86.00 |
| **This Research 1** | **GA + SVM** | **100** | **92.11** |
| **This Research 2** | **PSO + SVM** | **100** | **90.93** |

In comparison with the results of related research listed in Table 3.2, the proposed system achieves impressive performance. The PSO and GA-based feature selection methods account for these promising performances.

## 3.7 Chapter Summary

This research proposes an intelligent decision support system for the identification of benign and malignant skin lesions from dermoscopy images, which offers the prospect of achieving an improved and more accurate classification of lesions from images. The proposed system employs pre-processing techniques such as dull razors and median filters to remove hair and other noise. Then, the images are segmented using a pixel limitation technique to separate lesions from the image background. Feature extraction is subsequently conducted. The features extracted by the proposed system reflect the well-known dermatological parameters of ABCD as well as the criteria of the ELM. The features focus primarily on the size, shape, colour and local parameters of lesions with

some additional consideration given to the lesion edges. PSO and GA are also applied to identify the most discriminative feature subsets to improve classification accuracy. In an evaluation using 100 images from the Dermofit dataset, the system achieves average accuracy rates of 92.11% and 90.93% for GA and PSO respectively in skin lesion classification. Future research will concentrate on developing and refining a new hybrid feature selection algorithm, based on the integration of PSO or GA with other search algorithms such as harmony search, to identify more discriminative features and further improve classification accuracy.

# CHAPTER 4
# PROPOSED METHODOLOGY 2 WITH HLPSO-BASED FEATURE SELECTION

Feature selection is one of the most useful techniques for solving complex classification problems and enhancing performances. As indicated in the previous chapter, since the conventional PSO is lack of population and search diversity, a modified PSO model has been proposed to overcome premature convergence and local optimum traps. This chapter presents a newly developed optimisation method that is based on the fusion of PSO and GA operators to increase the exploration abilities of the PSO algorithm. The new model is referred to as hybrid learning PSO (HLPSO). Owing to the fact that the original PSO model does not include any mutation strategies, the proposed HLPSO model employs the GA operators, such as crossover and mutation, as well as the scatter operation to increase population diversity. A standard dataset has been chosen to assess the proposed algorithm. The capabilities of HLPSO are validated via comparison with numerous renowned standard optimization methods. This investigation proves that the proposed algorithm is a promising global optimisation algorithm and, moreover, that it is more advanced when compared with recently formulated versions of PSO in terms of classification performance and convergence speed.

Through the employment of a PSO-based feature selection system, this study offers a rational enhanced decision support system to achieve the classification of benign and malignant skin lesions. The system also incorporates four critical phases; namely, skin lesion segmentation, feature extraction, feature selection, and classification. Figure 4-1 provides an illustration of the system's configuration, indicating the main mechanisms of

the system, and the methodological approaches used are also explained in detail in this chapter.

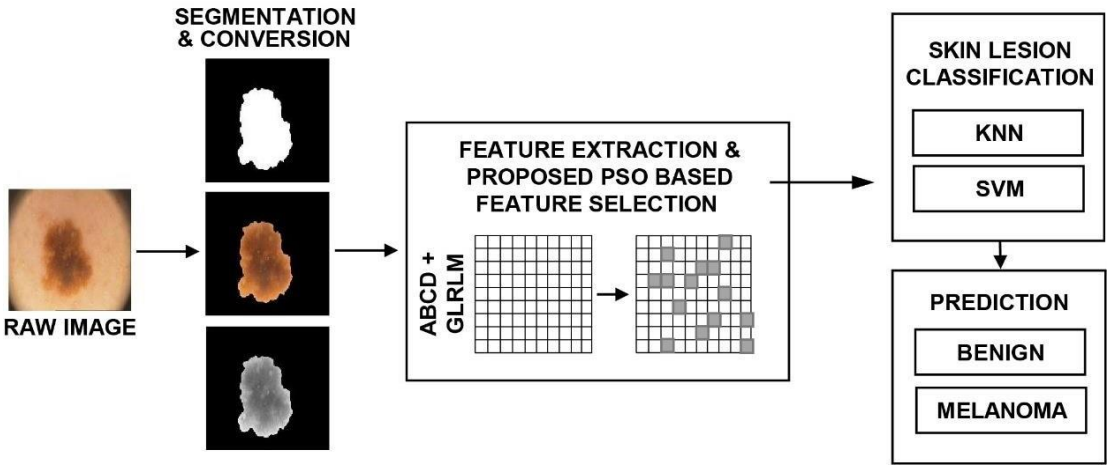



**Figure 4-1 : Skin Lesion Detection System**

## 4.1 Feature Extraction

The list of handcrafted features proposed in Chapter 1 has successfully achieved a promising level of accuracy. However, the speed and accuracy of feature selection could be improved further. The previous texture feature vector contains 3,888 features. Due to the overly dominance of the texture features than other features such as colour and shape, this has negative impact on classification accuracies. In this chapter, multiple new high-level feature extraction methods proposed in recent studies (Chang et al., 2013) are adopted, to produce colour, shape and Grey-Level Run Length Matrix (GLRLM) texture features for initial lesion representation. Furthermore, four orientation levels (i.e. 0°, 45°, 90°, and 135°) of the GLRLM-based texture features with eleven distinctive emphases (i.e. short-run emphasis, long-run emphasis, run percentage, run length nonuniformity, grey-level nonuniformity, low grey-level run emphasis, high grey-level run emphasis, short-run low grey-level emphasis, long-run low grey-level emphasis, short-run high grey-level emphasis and long-run high grey-level emphasis) are used for lesion representation. In addition, two extra texture features associated with the Tamura coarseness indices for skin and lesion are used as well. The chosen new features in this chapter are more related to the medical domain, which embed more colour attributes in multiple colour spaces to form a better classification performance as shown in Table 4.2.

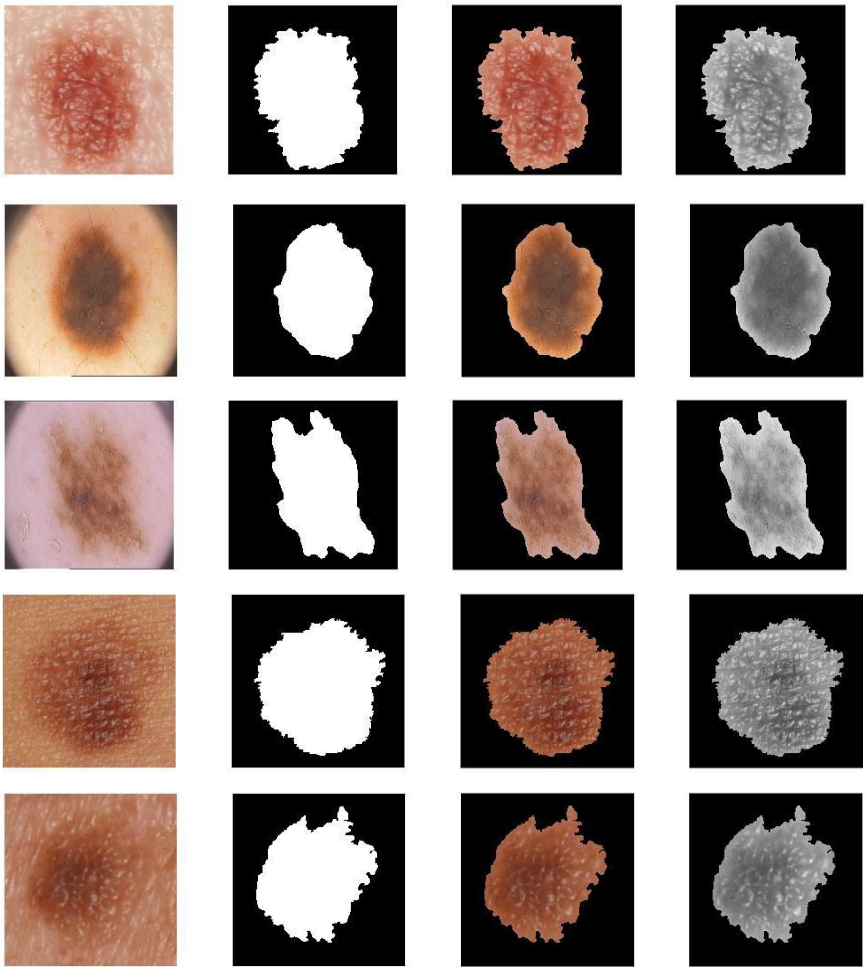


**Figure 4-2 : Raw Image (First column), Ground Truth (second column), Segmented Lesion (third column), and Grey Image (last column)**

**Table 4.1: The Extracted New Features**

| Features # | Features Details | Features Types | Extracted Area |
|---|---|---|---|
| 1 | Asymmetry index | Shape (13) | Lesion only |
| 2-4 | Compactness index | | |
| 5 | Radial variance | | |
| 6 | Border irregularity index | | |
| 7 | Perimeter | | |
| 8 | Solidity | | |
| 9 | Roundness | | |
| 10 | Extent | | |
| 11 | Equivalent diameter | | |
| 12 | Form factor | | |
| 13 | Different of left & right | | |
| 14-21 | Colour variance | Colour (87) | |
| 22-29 | Colour entropy | | |
| 30-37 | Colour skewness | | |
| 38-49 | Colour correlation | | |
| 50-55 | Principal component analysis variance | | |
| 56-58 | Mean of image darkness | | |

| | | | |
|---|---|---|---|
| 59-61 | Variance of image Darkness | | |
| 62-64 | Relative chromaticity | | |
| 65-76 | Mean and standard | | |
| 77-91 | Ratio of RGB | | |
| 92-100 | Average colour of RGB | | |
| 101-102 | Tamura coarseness index | | |
| 103-146 | Grey-level run length matrix with 4 orientations of GLRLM (0, 45, 90, 135), each level with 11 different emphasis | Texture (46) | Skin and Lesion |

A newly proposed feature set listed in Table 4.1, consisting of 146 handcrafted features (13 morphological, 87 colour and 46 texture features), is extracted to characterise the lesion area. Table 4.1 provides detailed information on these features. In addition, without applying any feature optimisation process, this feature set has been tested using KNN classifier and better accuracy rates have been achieved with significantly reduced cost as compared with those of the previous proposed feature set in Chapter 3. Table 4.2 shows that the newly proposed features outperformed the feature set used in the previous chapter.

**Table 4.2: The performance of the new features tested using the PH2 dataset before applying feature optimisation**

| | **Number of features** | **Accuracy (%)** |
|---|---|---|
| Previous Features Set used in Chapter 3 | 3914 | 76.00 |
| **Newly Proposed Features Set** | **146** | **85.20** |

## 4.2 The Proposed HLPSO for Feature Selection

In this chapter, the new PSO model (i.e. HLPSO) is proposed by integrating the sub-swarm division and crossover, mutation and scattering techniques with probability-based local exploration to moderate the local optimum trap. Algorithm 3 is shown below with a step-by-step demonstration of its operation. At first, the algorithm is initiated with 50 randomly chosen particles and then the initial fitness of each particle is calculated using fitness function. The algorithm uses the evaluated fitness for sorting, and the best particle is selected as the global best, $gbest$, which is stored in the best memory. The swarm is then divided into two sub-swarms.

Each sub-swarm will be led by different search methods, including probability distributions and conventional PSO operation respectively. In the first sub-swarm, three short- and long-jump exploration methods have been employed, i.e. Gaussian, Cauchy and Lévy distributions. Each particle in this sub-swarm will use one probability distribution method randomly for position updating. The new $gbest$ will then be identified by sorting the first sub-swarm. The second sub-swarm uses the conventional PSO algorithm for velocity and position updating with the identified best solution as the sub-swarm leader. Finally, the two sub-swarms are combined. The $gbest$ is then chosen by sorting the overall swarm based on the fitness scores. Top 10 and the bottom 20 particles are identified respectively. Subsequently, to improve the search diversity, 10 offspring will be produced using a crossover method by using the top 10 particles identified above. Furthermore, another set of 10 offspring particles will be generated by mutating the 10 newly generated particles from the crossover operation. In total, 20 offspring particles are generated from crossover and mutation operations. Then, it scatters the bottom 20 weak particles to random locations. Finally, the newly generated 20 offspring particles using mutation and crossover operations combine with the original top 10 and the scattered new 20 particles to form a new swarm, which is ranked based on the fitness scores. The algorithm continues until stagnation occurs.

**Algorithm 4.1: Pseudo-code of The Proposed HLPSO Model**

| | |
|---|---|
| **1** | **Start** |
| **2** | Initialize a population randomly (e.g. 50 particle); |
| **3** | Evaluate the population; |
| **4** | Sort population based on fitness values and identify $gBest$; |
| **5** | **While** (! Stop condition) do {// 358 iterations |
| **6** | Divide the population into two sub-swarms; |
| **7** | **For** (each particle in sub-swarm 1) do { |
| **8** | Conduct long jumps using Levy flights/Gaussian/Cauchy Eq. (4.3)-(4.5); |
| **9** | Update the $pBest$ if the new position developed by long jumps has better fitness; |
| **10** | } **End For** |
| **11** | **For** (each particle in sub-swarm 2) do { |
| **12** | Follow the leader in every dimension using Eq. (3.7) & (3.8); |
| **13** | Update the $pBest$ if the new position developed by the original PSO operation has better fitness; |
| **14** | } **End For** |
| **15** | Combine two sub-swarms; |
| **16** | Sort the overall swarm based on fitness values and identify the $gBest$; |

| | |
|---|---|
| **17** | Identify top 10 particles and scatter the last 20 particles to random locations; |
| **18** | Generate 10 offspring particles using crossover from the top 10 particles; |
| **19** | Generate another 10 offspring particles by mutating the above newly generated 10 offspring particles using the crossover operation; |
| **20** | Combine all 50 particles (top 10 + scattered 20 + 10 offspring generated using the crossover operation of the top 10 particles + 10 offspring generated by mutating the above 10 new particles produced using crossover); |
| **21** | Sort the population based on fitness value and identify the $gBest$; |
| **22** | } **End While** |
| **23** | **End** |

Due to imbalanced class samples used in the current experiment, the geometric mean (GM) is used as part of the fitness evaluation instead of classification accuracy, since the GM is often used to deal with imbalanced data problems (Beyan and Fisher, 2015, Srisukkham et al., 2017). Therefore, traditional performance measurements including accuracy and sensitivity will not be considered in this research due to the lack of equal treatment on each class or lack of efficiency. Since the GM measurement has been widely adopted in the field with the consideration of both sensitivity and specificity and it treats each class fairly without any bias on the sample numbers within a particular class, it is a more reliable performance measure. The GM and fitness evaluation formulas are shown below:

$$GM = \sqrt{Sensitivity * Specificity} \tag{4.1}$$

$$fitness(C) = (w_a * GM) + (w_f * ((ns - nf)/ns)) \tag{4.2}$$

where $w_a$ and $w_f$ represent the weights for classification accuracy and the number of selected features respectively. Again, a higher value is assigned to $w_a$ (e.g. 0.9) as opposed to $w_f$ (e.g. 0.1) owing to the importance of classification performance.

### 4.2.1 Gaussian Distribution

To reach the algorithm superior explorative capacity, several mutations are carried out to prevent premature convergence and local optima. The equation below permits the local exploitation of each particle in the first sub-swarm based on a Gaussian distribution (Jordehi, 2015).

$$np_g1(p) = cp_g1(p) + (X_{max} - X_{min}) * \varphi(o, h) \quad (4.3)$$

The Gaussian distribution is denoted by $\varphi(o,\ h)$, with $o$ and $h$ respectively denoting the distribution mean or expectation and the standard deviation, which exhibits a linear decline during execution. The upper bound of the decision vectors in the $dth$ dimension is $X_{max}$ while the lower bound is $X_{min}$ , with $d$ = 1, 2…$n$. Furthermore, the current particle, $cp_g1$, is replaced by the newly produced mutated particle, $np_g1$, if the new particle shows better fitness compared to that of the previous particle.

### 4.2.2 Cauchy Distribution

Cauchy mutation is also used to increase local exploitation of each particle and the position of each new particle is determined by the equation below (Jordehi, 2015):

$$np_g2(p) = cp_g2(p) + (X_{max} - X_{min}) * \psi(o, s) \quad (4.4)$$

The Cauchy mutation is denoted by $\psi(o,\ s)$, with $o$ denoting the location parameter, which shows where the peak is located, and the $s$ as scale parameter, which exhibits a linear decline during execution. If the fitness of $np_g2$ is greater than that of the previous $cp_g2$, then the better solution will replace the previous one.

### 4.2.3 Lévy Distribution

To further improve the exploration of each particle, the Lévy distribution is also used alongside the Gaussian distribution and the Cauchy distributions. The equation below illustrates Lévy's probability distribution (Lévy and Borel, 1954, Hakli and Uguz, 2013).

$$p_g(p) = p_g(p) + (X_{max} - X_{min}) * Levy(\mu, k, \eta) \quad (4.5)$$

The scale factor $\mu$ and the Lévy index have values in the range of [-1, 1]. Permitting adjustment, the parameter η is a key determinant shape of the distribution. For instance, smaller and larger η values allow the algorithm to undertake jumps of longer and shorter

lengths, respectively. Furthermore, a Cauchy distribution is applied when the value of η is 1 and a Gaussian distribution when it is 2. Hence, in contrast to both these types of distributions, a distribution with a more extended tail can be generated when η is greater than 0 but smaller than 1, enabling the performance of longer jumps (Lévy and Borel, 1954, Hakli and Uguz, 2013). Therefore, owing to the Lévy probability distribution, an offspring can jump a greater distance from the parent, allowing the detection of the best solutions in a wider search space.

## 4.3 Evaluation

The proposed PSO is evaluated through a comparison with other well-known optimisation algorithms. The state-of-the-art methods that have been implemented for comparison are the Bat Algorithm (BA) (Yang and Hossein Gandomi, 2012), Dragonfly Algorithm (DA) (Mirjalili, 2016), Harmony Search (HS) (Geem et al., 2001), Flower Pollination Algorithm (FPA) (Yang, 2012), Moth-Flame Optimisation (MFO) (Mirjalili, 2015), Artificial Bee Colony (ABC) (Hakli and Uguz, 2013), Cultural Algorithm (CA) (Reynolds, 1999), GA (Sivaraj and Ravichandran, 2011) and PSO (Kennedy, 2011).

As in the previous chapter, only one dataset from Dermofit has been used for evaluation. In this chapter, an additional dataset is used for the experimentation. That is the experiment combines two well-known datasets, which are the Dermatology Service of PH2, and Dermofit. The PH2 data includes 200 images with three categories of lesion images which include 80 benign, 80 atypical and 40 melanoma cancers. However, only 80 benign and 40 melanoma images were used as part of the proposed PSO evaluation. As introduced earlier, the Dermofit database has a total of 1,300 skin lesion images and includes ten categories of lesion images with 850 benign and 450 cancerous images. In this case, one of the benign skin types was utilised from this dataset for evaluation, which has 65 Dermatofibroma lesions and a cancer type with 76 malignant melanoma images in this dataset is also used. Therefore, a total of 145 benign lesion images with 80 from PH2 and 65 from Dermofit, and a total of 116 melanoma images with 40 from PH2 and 76 from Dermofit were selected from the two datasets for evaluation. Table 4.3 illustrates the number of images selected from each dataset.

**Table 4.3: Training and testing sets for skin cancer classification**

| Dataset | Benign | Melanoma |
|---|---|---|
| **PH2** | 80 | 40 |
| **Dermofit** | 65 | 76 |

This experiment utilised an aspect ratio of 80:20 for the training and testing sets respectively, which included 116 benign and 93 melanoma images for training, and 29 benign and 23 melanoma images for testing purposes. Thus, the training set contained 64 benign images from PH2 and 52 benign cases from Dermofit as well as 32 melanoma images from PH2 and 61 melanoma cases from Dermofit. Furthermore, the testing set contained 16 benign images from PH2 and 13 from Dermofit; it also included 8 melanoma images from PH2 and 15 from Dermofit. The reason of amending the dataset ratio from 70:30 to 80:20 is owing to the adoption of such an experimental setting in many related research studies, therefore, 80:20 is selected for performance comparison. In each case, the lesion is presented in the middle of the image with non-lesion skin visible at the corners of each image. Most of these images are presented in 8-bit RGB (red-green-blue) colour format in various resolutions. As most images incorporate various sizes of 300 - 600 pixels, resizing for all images was considered. This includes other images that did not qualify based on this requirement; they were edited to fit the required size of 500 pixels, while the same aspect ratio was sustained. Although the test datasets in this chapter are different from those used in the previous chapter, both PSO and GA algorithms from the previous chapter are used in experiments.

This experimentation utilised the proposed HLPSO method to determine the most relevant and reliable factors to conduct skin lesion classification. The GM was employed to compare the fitness outcomes with those generated from other methods. All feature selection algorithms subjected the extracted features to additional processing to reduce their dimensionality. Two types of skin lesion will be detected using the KNN and SVM classifiers, employing features produced automatically based on the feature selection method. The maximum numbers of fitness evaluation for every method were used as criteria for performance comparison. For instance, a termination criterion of, 50 population $\times$ 500 iterations $\times$ 1 function evaluation = 25,000 function evaluations, was employed in all search methods to provide meaningful and fair comparisons. For

example, during the feature selection process, a similar number of function evaluations, such as 70 population × 357 iterations × 1 function evaluation = 24,990 function evaluations, was employed by the HLPSO model, compared to 25,000 employed by other conventional approaches. This testing strategy was employed for all subsequent investigations to provide a meaningful comparison.

### 4.3.1 Parameter Settings

The best parameter settings for HLPSO and other optimisation methods were identified as shown in Table 4.4. In general, the parameter settings of the conventional PSO method are adopted for HLPSO, with some parameters being adjusted based on a trial-and-error approach. To determine how recurrent searches are affected by the prior speed, an inertia weight of 0.65 was employed. The acceleration constants for the HLPSO model were $c1 = C2 = 2.5$. For other search techniques, parameter settings are based on their original studies.

**Table 4.4: Parameter settings used in each method**

| Algorithms | Parameters |
|---|---|
| **BA** | Loudness = 0.5, pulse rate = 0.5 |
| **DA** | Separation factor = 0.1, alignment factor = 0.1, cohesion factor = 0.7, food factor = 1, enemy factor = 1, and inertia weight = 0.9 – m × ((0.9-0.4)/maxi_iterations), where m and maxi_iterations represent the current and maximum iteration numbers, respectively. |
| **HS** | Bandwidth = 0.2, harmony memory accepting rate = 0.95, pitch-adjusting rate = 0.3 |
| **FPA** | Switch/proximity probability = 0.8 |
| **MFO** | Use adaptive parameter settings |
| **ABC** | Food Sources = 50 (Population Size), abandonment limit = Round (0.6×dimensions×populations) |
| **CA** | Acceptance Ratio = 0.35, Number of Accepted Individuals = Round (Acceptance Ratio×Population Size) |
| **PSO** | Maximum Velocity = 0.6, Inertia Weight = 1.0, Acceleration Constants $C1 = C2 = 2$ |
| **HLPSO** | Maximum Velocity = 0.6, Inertia Weight = 0.60, Acceleration Constants $C1 = C2 = 2.5$ |

### 4.3.2 Evaluation Using the Combined Dataset

Initially, an evaluation using a combined dataset with both Dermofit and PH2 images was carried out by using 209 lesion images for training followed by 52 lesion images for testing. Table 4.6 presents the average classification performance for 30 runs with each optimisation algorithm combined with various classifiers. The KNN and SVM classifiers were implemented to evaluate the outcomes of each feature selection model. The KNN

model has achieved 91.98% average GM performance for 10-fold and 91.87% for hold-out validations. Our model outperformed all methods listed in Table 4.5. The result of KNN hold-out validation of our model also outperforms those of BA, DA, HS, FPA, MFO, ABC, CA, GA and PSO. Moreover, integrated with SVM, the average classification rates of HLPSO for two-lesion classification are 93.05% with 10-fold and 93.16% for hold-out, respectively, which are higher than the KNN classification results. The average result of SVM combined with HLPSO also outperformed those of the BA, DA, FPA, MFO, ABC, CA, GA and PSO. Lastly, for hold-out validation, our model outperformed the BA, FPA, MFO, ABC, CA, GA and PSO as shown table below:

**Table 4.5: Average performance correlation of each algorithm over 30 runs for the combined dataset**

| | KNN | | SVM | |
|---|---|---|---|---|
| | **10-fold** | **Hold-out** | **10-fold** | **Hold-out** |
| **BA** | 1.89% | 1.12% | 0.73% | 0.32% |
| **DA** | 2.06% | 1.15% | 0.39% | -0.32% |
| **HS** | 0.63% | 0.14% | -0.29% | -0.48% |
| **FPA** | 2.31% | 1.49% | 0.15% | 0.02% |
| **MFO** | 2.30% | 1.45% | 2.58% | 2.14% |
| **ABC** | 2.18% | 1.52% | 0.97% | 0.72% |
| **CA** | 1.61% | 1.02% | 0.37% | 0.09% |
| **GA** | 3.90% | 3.13% | 4.87% | 4.41% |
| **PSO** | 2.40% | 1.66% | 0.70% | 0.35% |

The comparisons above illustrate the superiority of the proposed HLPSO model, whose performance surpasses those of most of the other algorithms listed in Table 4.6; the only exception is DA and HS, which outperform the HLPSO model when combined with the SVM classifier under 10-fold and hold-out validations, respectively.

**Table 4.6: Average performance of each algorithm over 30 runs for the combined dataset**

| | KNN | | SVM | |
|---|---|---|---|---|
| | **10-fold** | **Hold-out** | **10-fold** | **Hold-out** |
| **BA** | 0.9009 | 0.9075 | 0.9232 | 0.9284 |
| **DA** | 0.8992 | 0.9072 | 0.9266 | 0.9348 |
| **HS** | 0.9135 | 0.9173 | **0.9334** | **0.9364** |
| **FPA** | 0.8967 | 0.9038 | 0.9290 | 0.9314 |
| **MFO** | 0.8968 | 0.9042 | 0.9047 | 0.9102 |
| **ABC** | 0.8980 | 0.9035 | 0.9208 | 0.9244 |
| **CA** | 0.9037 | 0.9085 | 0.9268 | 0.9307 |
| **GA** | 0.8808 | 0.8874 | 0.8818 | 0.8875 |
| **PSO** | 0.8958 | 0.9021 | 0.9235 | 0.9281 |
| **HLPSO** | **0.9198** | **0.9187** | 0.9305 | 0.9316 |

A statistical Wilcoxon rank sum test (Wilcoxon, 1945, Derrac et al., 2011) was used as additional evidence in the comparison, the results of which also show the advantage of the proposed HLPSO model. This is a non-parametric test which assesses whether the distributions of two populations, including those of the proposed HLPSO model and those from another method, possess a statistically equal median score. The rank sum test provides a *p-value* which may lead to a recommendation that the null hypothesis of an equal median should be rejected at less than 0.05 or otherwise. Table 4.7 shows the test outcomes for each algorithm, showing that almost all *p-values* are below 0.05. The exceptions are for the HS and FPA, where the HLPSO model has the same mean results as those of HS combined with KNN and FPA integrated with SVM, over 30 runs. Also, HS outperforms HLPSO when combined with SVM statistically.

**Table 4.7: The *p-values* of the Wilcoxon rank sum test using the combined dataset**

| | BA | DA | HS | FPA | MFO | ABC | CA | GA | PSO |
|---|---|---|---|---|---|---|---|---|---|
| **KNN Wilcoxon 10F** | 4.64E-05 | 4.03E-03 | **6.15E-02** | 2.77E-05 | 1.76E-02 | 2.53E-04 | 4.24E-02 | 1.49E-04 | 2.68E-05 |
| **KNN Wilcoxon HO** | 3.15E-05 | 1.76E-02 | **5.24E-02** | 4.03E-03 | 3.92E-02 | 2.77E-05 | 1.76E-02 | 2.53E-04 | 6.78E-03 |
| **SVM Wilcoxon 10F** | 2.68E-03 | 4.24E-02 | 1.49E-04 | **7.73E-02** | 6.15E-05 | 1.29E-02 | 8.88E-03 | 4.51E-08 | 5.10E-03 |
| **SVM Wilcoxon HO** | 7.22E-04 | 3.57E-02 | 2.15E-02 | **6.49E-01** | 1.78E-04 | 1.88E-04 | 3.15E-02 | 3.76E-06 | 3.78E-02 |

**Table 4.8: The training computational cost and the average number of features selected using the combined dataset**

| | HLPSO | BA | DA | HS | FPA | MFO | ABC | CA | GA | PSO |
|---|---|---|---|---|---|---|---|---|---|---|
| **Average no. of selected features** | 61.78 | 74.35 | 78.09 | 71.86 | 70.73 | 65.93 | 70.78 | 67.38 | 62.69 | 73.58 |
| **Training cost (seconds)** | 3807.72 | 3816.57 | 3813.44 | 3931.89 | 3806.12 | 3905.52 | 3906.70 | 3929.63 | 3791.70 | 3878.16 |

### 4.3.3 Evaluation Using the PH2 Dataset

Secondly, another dataset, i.e. PH2, has been used for evaluation, which contains 200 images including 80 benign, 80 atypical and 40 melanoma lesions. All three classes were used for evaluation and the images were separated into 80:20 aspect ratio for training and testing respectively. This includes 160 images for training, with 64 benign, 64 atypical and 32 melanoma images. A total of 40 images were used in the testing set with 16 benign, 16 atypical and 8 melanoma images. The average classification performances, which combined with various classifiers including SVM and KNN over 30 runs, are presented in Table 4.10. In this section, the average GM performances of the proposed HLPSO model for three-class lesion categorization achieved 93.88% in 10-fold and 95.62% for hold-out; the HLPSO model has been proven to be more efficient when compared with

other algorithms. It outperforms BA, DA, HS, FPA, MFO, ABC, CA, GA and PSO respectively for 10-fold validation, while for the hold-out validation, it outperformed the BA, DA, HS, FPA, MFO, ABC, CA, GA and PSO, respectively. When it comes to the SVM implementation, the average GM performances are 93.54% for 10-fold and 96.00% for hold-out validations. HLPSO outperformed the BA, DA, HS, MFO, ABC, CA, GA and PSO respectively for 10-fold validation and outperformed the BA, DA, HS, MFO, ABC, CA, GA and PSO, respectively, for hold-out validation.

**Table 4.9: Average performance correlation of each algorithm over 30 runs for the PH2 dataset**

| | KNN | | SVM | |
|---|---|---|---|---|
| | **10-fold** | **Hold-out** | **10-fold** | **Hold-out** |
| **BA** | 4.43% | 3.75% | 2.43% | 1.44% |
| **DA** | 5.43% | 3.53% | 1.88% | 1.83% |
| **HS** | 6.35% | 6.33% | 0.10% | 0.60% |
| **FPA** | 5.18% | 3.15% | -0.78% | -0.14% |
| **MFO** | 0.46% | 0.38% | 0.38% | 0.38% |
| **ABC** | 2.93% | 2.72% | 1.16% | 0.26% |
| **CA** | 3.38% | 3.06% | 0.14% | 0.17% |
| **GA** | 1.88% | 1.98% | 1.33% | 1.15% |
| **PSO** | 5.45% | 3.45% | 2.61% | 1.32% |

The comparison above again illustrates the superiority of the proposed HLPSO model, whose performance surpasses those of most of the other algorithms. The only exception here is FPA, which outperforms the HLPSO when combined with the SVM classifier under 10-fold and hold-out validations. In other cases, algorithms such as HS, FPA, MFO and CA show similar result distributions in comparison with those of the HLPSO model when integrated with KNN classification under hold-out validation, and SVM classifier under 10-fold and hold-out validations, respectively. In general, the outcomes of the proposed algorithm are statistically better than the outcomes from other methods in almost all test cases.

**Table 4.10: Average performance of each algorithm over 30 runs for the PH2 dataset**

| | KNN | | SVM | |
|---|---|---|---|---|
| | **10-fold** | **Hold-out** | **10-fold** | **Hold-out** |
| **BA** | 0.8945 | 0.9187 | 0.9111 | 0.9456 |
| **DA** | 0.8845 | 0.9209 | 0.9166 | 0.9417 |
| **HS** | 0.8753 | 0.8929 | 0.9344 | 0.9540 |
| **FPA** | 0.8870 | 0.9247 | **0.9432** | **0.9614** |
| **MFO** | 0.9342 | 0.9524 | 0.9316 | 0.9562 |
| **ABC** | 0.9095 | 0.9290 | 0.9238 | 0.9574 |
| **CA** | 0.9050 | 0.9256 | 0.9340 | 0.9583 |
| **GA** | 0.9200 | 0.9364 | 0.9221 | 0.9485 |
| **PSO** | 0.8843 | 0.9217 | 0.9093 | 0.9468 |
| **HLPSO** | **0.9388** | **0.9562** | 0.9354 | 0.9600 |

**Table 4.11: The *p-values* of the Wilcoxon rank sum test using PH2 dataset**

| | **BA** | **DA** | **HS** | **FPA** | **MFO** | **ABC** | **CA** | **GA** | **PSO** |
|---|---|---|---|---|---|---|---|---|---|
| **KNN Wilcoxon 10F** | 2.21E-06 | 1.32E-08 | 2.18E-06 | 1.91E-07 | 1.89E-02 | 1.33E-08 | 1.33E-08 | 2.04E-05 | 7.43E-10 |
| **KNN Wilcoxon HO** | 1.04E-06 | 1.57E-05 | 4.32E-07 | 1.80E-07 | **6.19E-02** | 6.91E-07 | 1.30E-06 | 1.43E-05 | 9.44E-08 |
| **SVM Wilcoxon 10F** | 1.80E-02 | 8.37E-04 | **3.46E-01** | 1.24E-03 | 1.80E-02 | 4.32E-04 | **6.40E-01** | 1.04E-06 | 1.89E-07 |
| **SVM Wilcoxon HO** | 4.38E-08 | 2.65E-08 | 5.25E-03 | **2.64E-01** | 1.24E-03 | 6.98E-06 | 1.01E-03 | 1.48E-05 | 7.14E-06 |

**Table 4.12: The training computational cost and the average number of features selected using PH2 dataset**

| | **HLPSO** | **BA** | **DA** | **HS** | **FPA** | **MFO** | **ABC** | **CA** | **GA** | **PSO** |
|---|---|---|---|---|---|---|---|---|---|---|
| **Average no. of selected features** | 60.22 | 68.60 | 69.60 | 65.33 | 64.83 | 57.70 | 66.64 | 61.37 | 62.20 | 66.63 |
| **Training cost (seconds)** | 3772.39 | 3808.29 | 3806.97 | 3927.84 | 3798.13 | 3900.98 | 3897.18 | 3922.39 | 3782.29 | 3876.96 |

### 4.3.4 Evaluation Using Benchmark Functions

We also use benchmark functions, i.e. unimodal and multimodal functions, for the evaluation of the proposed HLPSO model. Such functions are particularly useful in this respect due to their varied shapes and other challenging characteristics. The evaluation of the proposed PSO and other algorithms was conducted by applying a series of widely used benchmark functions including Dixon Price, Sphere, Rotated Hyper-Ellipsoid, Sum Squares, Sum of Different Powers, Ackley, Griewank and Powell. In theoretical terms, diversity in benchmark function properties is essential, because a limited sample can introduce bias into the outcomes. Each algorithm was conducted with an identical number of function evaluations. Each experiment was performed with 30 runs for assessments. In each of these functions, the real global optimum is zero. The experiments were performed in three dimensional settings, i.e. low (e.g. 10), medium (e.g. 30) and high (e.g. 50) dimensions, to generate a comprehensive assessment of how the proposed PSO performs under various scenarios. A statistical appraisal of the attained results was undertaken in a

pair-wise manner using the Wilcoxon rank sum test. It is important to note that the approach implemented by (Zhang et al., 2018a) has been taken as a model regarding which test functions to use in this study. A summary of the relevant benchmark functions experiment can be found in Table 4.13, where $S$ represents the variable scales, and $Fmin$ represents the optimum global value in $S$.

**Table 4.13: Unimodal and multimodal benchmark functions**

| | Function Name | Function | S | *Fmin* |
|---|---|---|---|---|
| **F1** | Dixon-Price | $f(x) = (x_1 - 1)^2 + \sum_{i=2}^{d} i(2x_i^2 - x_{i-1})^2$ | (-10, 10) | 0 |
| **F2** | Sphere | $f(x) = \sum_{i=1}^{d} x_i^2$ | (-100, 100) | 0 |
| **F3** | Rotated Hyper-Ellipsoid | $f(x) = \sum_{i=1}^{d} \sum_{j=1}^{i} x_j^2$ | (-65.536, 65.536) | 0 |
| **F4** | Sum Squares | $f(x) = \sum_{i=1}^{d} ix_i^2$ | (-5.12, 5.12) | 0 |
| **F5** | Sum of Different Powers | $f(x) = \sum_{i=1}^{d} \lvert x_i \rvert^{i+1}$ | (1-, 1) | 0 |
| **F6** | Ackley | $f(x) = -a\,exp\left(-b\sqrt{\frac{1}{d}\sum_{i=1}^{d} x_i^2}\right) - \exp\left(\frac{1}{d}\sum_{i=1}^{d}\cos(cx_i)\right)$ | (-32, 32) | 0 |
| **F7** | Griewank | $f(x) = \sum_{i=1}^{d} \frac{-x_i^2}{4000} - \prod_{i=1}^{d} \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1$ | (-600, 600) | 0 |
| **F8** | Powell | $f(x) = \sum_{i=1}^{d}[(x_{4i-3} + 10x_{4i-2})^2 + 5(x_{4i-1} + x_{4i})^2 + (x_{4i-2} + 2x_{4i-1})^4 + 10(x_{4i-3} + 10x_{4i})^4]$ | (-4, 5) | 0 |

It is important to recognise that, for specific applications, identifying an effective optimal or suboptimal solution is crucial. Meanwhile, when contrasting applications, the precision of the solution is the guiding principle. As indicated in Table 4.13, the F1 - F5 functions are unimodal landscapes. On the other hand, the F6 - F8 functions are multimodal functions. The F6 Ackley function is a commonly employed multimodal benchmark function that has been incorporated into several studies and is widely used (Zhang et al., 2018a). Particularly, F6 is also a typically difficult problem to solve. It is typified by an extremely deep-set and centrally-located valley surrounded by a relatively level area, and this is what causes the difficulty in generating an adequate solution. In almost all cases, optimisation algorithms can be rapidly pulled into local minima.

In each evaluation, the number of iterations is adjusted by ensuring that every algorithm is measured based on the same function evaluation frequencies of 500 iterations × 50 populations × 1 function evaluation = 25,000 function evaluation. This guarantees fairness in the comparative experimentation against all other methods. Therefore, the maximum iteration frequency for the HLPSO algorithm is established at 70 population × 357 iterations × 1 function evaluation = 24,990 function evaluation, and the numbers of dimensions established for benchmark functions are equivalent to 10, 30, and 50. Additionally, 30 runs with distinct initial configurations were conducted on every function and algorithm. The algorithm outcomes are accompanied by several conventional statistical distributions including mean, standard deviation (std), and minimum (min) and maximum (max) of the fitness score generated with 30 runs. Table 4.14 - Table 4.19 present tabulations of the statistical values resulting from the analysis of unimodal and multimodal benchmark test functions, with the optimal figure for every row emboldened in red. The mean of the fitness values is the most important point when comparing the algorithms, and it is particularly notable that the proposed HLPSO model has a powerful exploitation ability.

### I. Results for Benchmark Functions with Low Dimension ($D = 10$)

When compared to other optimisation algorithms for unimodal problems with a dimension 10, the proposed PSO model obtains the closet results to the global optima. It performs more efficiently than the alternative optimisation algorithms for functions F1 to F8. The key finding from Table 4.14 is that the proposed PSO strategic approach, based on successive sub-swarm, crossover and mutation methods, results in a highly efficient exploration of the search space, thereby having a considerably mitigatory impact on the premature convergence. Figure 4.3 presents the convergence curves for all algorithms with specific functions. The *p-values* derived from the Wilcoxon rank sum test for dimension 10 are presented in a formation form in Table 4.15, which demonstrates that HLPSO generates optimistic outcomes when compared to almost every optimisation algorithm. Although HLPSO performs more efficiently than MFO and CA, the *p-value* of Dixon-Price and Ackley is not satisfactorily small, whereas most of the other *p-values* are lower than 0.05. Overall, HLPSO outperforms other methods in most of the test cases.

Figure 4.3 indicates that HLPSO shows the fastest convergence rate in comparison with those of other methods.

**Table 4.14 : Evaluation results for benchmark functions with $D$ = 10**

| | | HLPSO | BA | DA | HS | FPA | MFO | ABC | CA | GA | PSO |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **F1** | **Mean** | **4.89E-01** | 4.55E+01 | 9.61E+00 | 6.14E-01 | 4.85E+00 | 3.56E+01 | 6.67E-01 | 7.17E-01 | 1.78E+03 | 5.31E+03 |
| | **Std** | 3.00E-01 | 7.56E+01 | 2.87E+01 | 3.35E-01 | 2.01E+00 | 8.36E+01 | **1.74E-03** | 1.91E-01 | 2.15E+03 | 3.35E+03 |
| | **Min** | **5.29E-20** | 6.67E-01 | 6.67E-01 | 1.80E-02 | 1.96E+00 | 3.64E-09 | 6.67E-01 | 3.38E-03 | 8.27E+01 | 8.28E+02 |
| | **Max** | **6.67E-01** | 3.42E+02 | 1.24E+02 | 1.01E+00 | 9.30E+00 | 3.21E+02 | 6.76E-01 | 1.18E+00 | 9.76E+03 | 1.59E+04 |
| **F2** | **Mean** | **1.49E-35** | 1.05E-06 | 3.09E-03 | 1.05E-04 | 1.26E-01 | 2.47E-16 | 2.90E-11 | 2.50E-28 | 1.71E+00 | 1.11E+01 |
| | **Std** | **4.92E-35** | 2.19E-07 | 5.27E-03 | 5.05E-05 | 5.60E-02 | 4.10E-16 | 3.86E-11 | 1.21E-27 | 1.13E+00 | 3.25E+00 |
| | **Min** | 4.47E-42 | 6.37E-07 | **0.00E+00** | 2.87E-05 | 4.35E-02 | 6.67E-19 | 1.56E-12 | 1.33E-52 | 3.96E-01 | 3.54E+00 |
| | **Max** | **2.56E-34** | 1.61E-06 | 1.85E-02 | 2.69E-04 | 3.24E-01 | 1.54E-15 | 1.53E-10 | 6.62E-27 | 5.52E+00 | 1.83E+01 |
| **F3** | **Mean** | **2.81E-33** | 5.21E+03 | 2.82E+01 | 5.83E-04 | 8.69E+01 | 1.43E+02 | 1.30E-08 | 7.73E-29 | 3.54E+03 | 9.06E+03 |
| | **Std** | **5.93E-33** | 2.27E+03 | 8.02E+01 | 2.65E-04 | 3.23E+01 | 7.84E+02 | 1.28E-08 | 4.23E-28 | 2.88E+03 | 3.20E+03 |
| | **Min** | 3.82E-39 | 1.76E+03 | 1.08E-30 | 1.34E-04 | 9.08E+00 | 2.49E-16 | 1.01E-09 | **9.14E-49** | 9.99E+02 | 2.12E+03 |
| | **Max** | **2.46E-32** | 1.04E+04 | 4.05E+02 | 1.26E-03 | 1.34E+02 | 4.29E+03 | 5.43E-08 | 2.32E-27 | 1.61E+04 | 1.71E+04 |
| **F4** | **Mean** | **2.57E-35** | 8.16E+00 | 3.31E-01 | 4.28E-04 | 2.33E+00 | 2.82E-15 | 5.51E-10 | 8.26E-35 | 4.87E+01 | 1.79E+02 |
| | **Std** | **7.78E-35** | 1.15E+01 | 5.33E-01 | 1.73E-04 | 9.11E-01 | 4.56E-15 | 1.07E-09 | 4.45E-34 | 2.29E+01 | 6.45E+01 |
| | **Min** | 8.00E-40 | 6.18E-06 | **0.00E+00** | 1.73E-04 | 9.95E-01 | 9.50E-18 | 1.79E-11 | 2.71E-51 | 1.82E+01 | 6.76E+01 |
| | **Max** | **4.13E-34** | 4.84E+01 | 2.67E+00 | 7.78E-04 | 5.47E+00 | 1.58E-14 | 4.42E-09 | 2.44E-33 | 9.76E+01 | 3.26E+02 |
| **F5** | **Mean** | **1.79E-65** | 1.33E-08 | 4.33E-06 | 3.56E-08 | 4.77E-06 | 8.10E-35 | 1.66E-17 | 9.48E-11 | 1.03E-02 | 1.88E-02 |
| | **Std** | **9.80E-65** | 6.23E-09 | 9.72E-06 | 4.84E-08 | 4.03E-06 | 3.42E-34 | 3.16E-17 | 2.96E-10 | 1.46E-02 | 1.02E-02 |
| | **Min** | **1.77E-85** | 1.33E-09 | 2.73E-16 | 7.82E-12 | 5.54E-07 | 3.32E-42 | 4.02E-20 | 1.60E-36 | 3.39E-05 | 2.93E-03 |
| | **Max** | **5.37E-64** | 2.54E-08 | 5.17E-05 | 2.05E-07 | 1.89E-05 | 1.86E-33 | 1.43E-16 | 1.27E-09 | 6.26E-02 | 4.08E-02 |
| **F6** | **Mean** | **7.40E-15** | 1.38E+01 | 2.02E+00 | 2.08E-02 | 6.46E+00 | 4.87E-08 | 6.21E-04 | 3.48E-01 | 1.13E+01 | 1.64E+01 |
| | **Std** | **3.11E-15** | 1.97E+00 | 1.18E+00 | 7.78E-03 | 8.34E-01 | 3.98E-08 | 3.68E-04 | 7.17E-01 | 1.90E+00 | 1.22E+00 |
| | **Min** | **4.44E-15** | 9.00E+00 | **4.44E-15** | 8.30E-03 | 4.35E+00 | 4.63E-09 | 1.67E-04 | **4.44E-15** | 6.99E+00 | 1.34E+01 |
| | **Max** | **1.51E-14** | 1.67E+01 | 5.38E+00 | 4.19E-02 | 7.77E+00 | 1.74E-07 | 1.68E-03 | 3.03E+00 | 1.44E+01 | 1.81E+01 |
| **F7** | **Mean** | **5.62E-02** | 2.89E+01 | 4.19E-01 | 1.00E-01 | 1.38E+00 | 1.72E-01 | 3.11E-01 | 1.49E-01 | 1.74E+01 | 4.10E+01 |
| | **Std** | **2.43E-02** | 1.57E+01 | 3.62E-01 | 5.37E-02 | 1.32E-01 | 9.40E-02 | 6.43E-02 | 7.58E-02 | 8.02E+00 | 9.48E+00 |
| | **Min** | 1.72E-02 | 1.23E+01 | **0.00E+00** | 7.43E-03 | 1.07E+00 | 3.94E-02 | 1.64E-01 | 5.17E-02 | 3.80E+00 | 2.26E+01 |
| | **Max** | **1.28E-01** | 7.14E+01 | 1.57E+00 | 2.34E-01 | 1.60E+00 | 4.50E-01 | 4.66E-01 | 3.45E-01 | 3.32E+01 | 5.36E+01 |
| **F8** | **Mean** | **5.89E-07** | 1.14E-01 | 1.01E+01 | 8.04E-03 | 2.53E-02 | 2.12E+00 | 2.14E-04 | 7.94E-03 | 3.92E+01 | 1.08E+02 |
| | **Std** | **8.07E-07** | 5.26E-01 | 2.41E+01 | 4.37E-03 | 1.24E-02 | 1.16E+01 | 1.28E-04 | 2.09E-02 | 2.89E+01 | 4.06E+01 |
| | **Min** | **1.16E-08** | 1.27E-04 | 3.64E-03 | 2.85E-04 | 6.48E-03 | 1.95E-05 | 5.73E-05 | 2.17E-05 | 8.18E+00 | 4.30E+01 |
| | **Max** | **3.01E-06** | 2.88E+00 | 7.73E+01 | 1.99E-02 | 5.85E-02 | 6.36E+01 | 5.20E-04 | 1.14E-01 | 1.19E+02 | 2.26E+02 |

**Table 4.15: The Wilcoxon rank sum test results for all benchmark functions with $D$ = 10**

| | BA | DA | HS | FPA | MFO | ABC | CA | GA | PSO |
|---|---|---|---|---|---|---|---|---|---|
| **F1** | 8.08E-11 | 2.99E-11 | 5.25E-05 | 2.99E-11 | **8.76E-02** | 1.94E-10 | **1.19E-01** | 2.99E-11 | 2.99E-11 |
| **F2** | 3.02E-11 | 8.48E-09 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 8.48E-09 | 3.02E-11 | 3.02E-11 |
| **F3** | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 1.17E-09 | 3.02E-11 | 3.02E-11 |
| **F4** | 3.02E-11 | 8.48E-09 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 1.31E-08 | 3.02E-11 | 3.02E-11 |
| **F5** | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 |
| **F6** | 1.69E-11 | 1.89E-10 | 1.69E-11 | 1.69E-11 | 1.69E-11 | 1.69E-11 | **2.39E-01** | 1.69E-11 | 1.69E-11 |
| **F7** | 3.02E-11 | 1.36E-07 | 3.18E-04 | 3.02E-11 | 5.00E-09 | 3.02E-11 | 7.12E-09 | 3.02E-11 | 3.02E-11 |
| **F8** | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 |

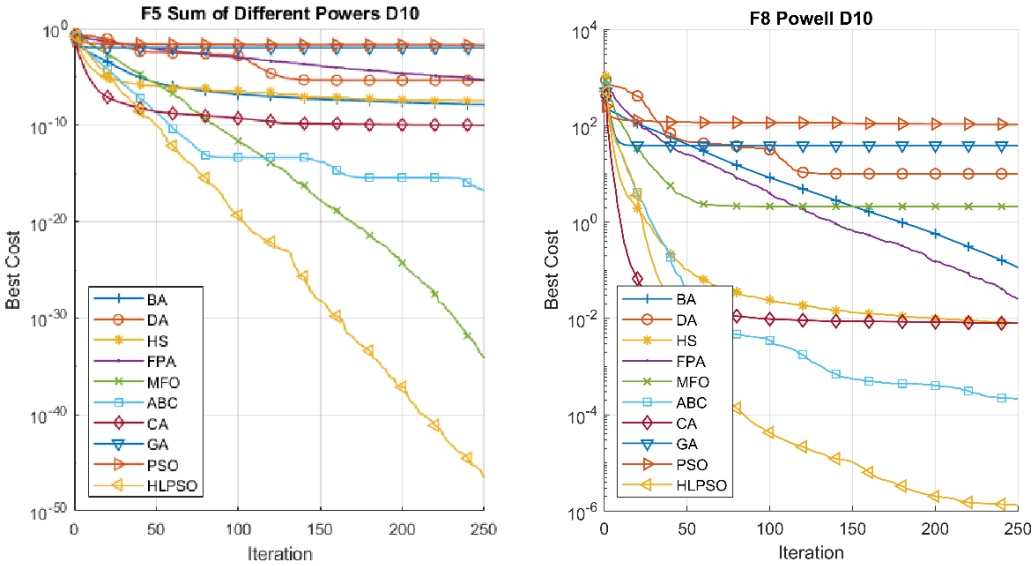



**Figure 4.3: Convergence curves for F5 and F8 in *D* = 10**

## II. Results for Benchmark Functions with Medium Dimension (*D* = 30)

For the medium dimension, every algorithm was established with an identical parameter, and Table 4.16 presents the formation of the outcomes for each algorithm regarding the unimodal and multimodal functions. Evidently, the proposed method outperformed the other optimisation algorithms. The exploration ability of the proposed PSO is reflected in Figure 4-4, with the convergence curves of F5 and F7. The *p-values* resulting from the Wilcoxon rank sum test for this dimension are presented in Table 4.17; the table shows that the HLPSO generates optimistic outcomes when compared to every optimisation algorithm. All the *p-values* are lower than 0.05; this indicates that the HLPSO model has achieved statistical improvements over all other methods.

**Table 4.16 : Evaluation results for benchmark functions with *D* = 30**

| | | HLPSO | BA | DA | HS | FPA | MFO | ABC | CA | GA | PSO |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **F1** | **Mean** | **1.03E+00** | 1.58E+04 | 4.18E+02 | 4.86E+00 | 5.50E+03 | 4.37E+04 | 1.58E+03 | 6.85E+01 | 7.42E+04 | 1.58E+04 |
| | **Std** | **7.34E-01** | 1.71E+04 | 8.03E+02 | 1.78E+00 | 2.66E+03 | 8.99E+04 | 7.83E+02 | 9.79E+01 | 4.23E+04 | 1.71E+04 |
| | **Min** | **6.67E-01** | 2.88E+02 | 1.21E+01 | 2.51E+00 | 2.17E+03 | 3.70E+00 | 6.40E+02 | 6.58E+00 | 1.95E+04 | 2.88E+02 |
| | **Max** | **3.30E+00** | 7.10E+04 | 4.43E+03 | 9.08E+00 | 1.26E+04 | 3.46E+05 | 3.81E+03 | 4.73E+02 | 1.54E+05 | 7.10E+04 |
| **F2** | **Mean** | **1.65E-10** | 2.43E+00 | 2.56E+00 | 6.00E-02 | 7.79E+00 | 4.37E+00 | 2.65E-02 | 7.03E-03 | 1.69E+01 | 2.43E+00 |
| | **Std** | **2.06E-10** | 2.41E+00 | 2.19E+00 | 7.48E-03 | 2.33E+00 | 9.94E+00 | 8.87E-03 | 1.04E-02 | 8.55E+00 | 2.41E+00 |
| | **Min** | **4.64E-12** | 1.61E-05 | 3.14E-01 | 5.00E-02 | 2.29E+00 | 5.48E-04 | 1.19E-02 | 9.21E-09 | 6.95E+00 | 1.61E-05 |
| | **Max** | **1.05E-09** | 8.73E+00 | 7.81E+00 | 7.95E-02 | 1.39E+01 | 2.62E+01 | 4.97E-02 | 4.74E-02 | 4.81E+01 | 8.73E+00 |
| **F3** | **Mean** | **2.34E-08** | 8.34E+04 | 5.08E+03 | 2.06E+00 | 1.66E+04 | 1.93E+04 | 4.14E+01 | 2.23E+01 | 7.43E+04 | 8.34E+04 |
| | **Std** | **2.38E-08** | 2.82E+04 | 2.73E+03 | 5.83E-01 | 4.06E+03 | 1.89E+04 | 1.19E+01 | 4.68E+01 | 2.26E+04 | 2.82E+04 |
| | **Min** | **2.54E-09** | 3.67E+04 | 6.71E+02 | 1.38E+00 | 8.93E+03 | 1.17E+00 | 2.35E+01 | 3.51E-03 | 4.18E+04 | 3.67E+04 |
| | **Max** | **1.13E-07** | 1.41E+05 | 1.26E+04 | 4.48E+00 | 2.37E+04 | 6.44E+04 | 7.51E+01 | 2.20E+02 | 1.37E+05 | 1.41E+05 |
| **F4** | **Mean** | **5.83E-09** | 6.71E+02 | 1.25E+02 | 9.39E-01 | 3.75E+02 | 5.10E+02 | 1.10E+00 | 1.37E+00 | 1.11E+03 | 6.71E+02 |
| | **Std** | **7.90E-09** | 3.30E+02 | 1.13E+02 | 1.57E-01 | 9.80E+01 | 6.17E+02 | 2.93E-01 | 3.11E+00 | 3.50E+02 | 3.30E+02 |
| | **Min** | **1.11E-10** | 7.65E+01 | 1.80E+00 | 5.66E-01 | 1.87E+02 | 4.40E-02 | 6.01E-01 | 2.38E-03 | 2.91E+02 | 7.65E+01 |
| | **Max** | **3.77E-08** | 1.75E+03 | 4.88E+02 | 1.18E+00 | 6.06E+02 | 2.40E+03 | 1.82E+00 | 1.26E+01 | 1.83E+03 | 1.75E+03 |
| **F5** | **Mean** | **3.59E-29** | 6.43E-08 | 1.18E-05 | 2.41E-07 | 1.62E-04 | 3.07E-11 | 4.10E-03 | 5.20E-07 | 1.93E-02 | 6.43E-08 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Std** | **1.01E-28** | 3.20E-08 | 2.42E-05 | 3.15E-07 | 1.60E-04 | 6.37E-11 | 3.17E-03 | 1.25E-06 | 2.21E-02 | 3.20E-08 |
| | **Min** | **1.54E-35** | 1.76E-08 | 3.74E-10 | 1.10E-09 | 1.10E-05 | 1.02E-14 | 2.63E-04 | 7.67E-10 | 8.19E-04 | 1.76E-08 |
| | **Max** | **4.28E-28** | 1.73E-07 | 1.32E-04 | 1.48E-06 | 6.97E-04 | 2.51E-10 | 1.52E-02 | 6.35E-06 | 8.77E-02 | 1.73E-07 |
| **F6** | **Mean** | **2.03E-04** | 1.52E+01 | 7.76E+00 | 2.18E+00 | 1.31E+01 | 4.77E+00 | 3.85E+00 | 4.19E+00 | 1.45E+01 | 1.52E+01 |
| | **Std** | **7.77E-04** | 1.16E+00 | 1.91E+00 | 4.68E-01 | 9.53E-01 | 6.75E+00 | 8.62E-01 | 1.23E+00 | 1.19E+00 | 1.16E+00 |
| | **Min** | **2.38E-06** | 1.28E+01 | 2.06E+00 | 6.00E-01 | 1.07E+01 | 2.20E-01 | 2.64E+00 | 2.32E+00 | 1.25E+01 | 1.28E+01 |
| | **Max** | **4.26E-03** | 1.71E+01 | 1.09E+01 | 2.79E+00 | 1.47E+01 | 1.91E+01 | 6.65E+00 | 6.68E+00 | 1.69E+01 | 1.71E+01 |
| **F7** | **Mean** | **2.23E-02** | 1.50E+02 | 1.02E+01 | 1.09E+00 | 2.95E+01 | 2.49E+01 | 1.07E+00 | 4.91E-01 | 1.11E+02 | 1.50E+02 |
| | **Std** | **3.14E-02** | 5.20E+01 | 7.45E+00 | 3.66E-02 | 8.26E+00 | 4.69E+01 | 3.72E-02 | 4.91E-01 | 3.36E+01 | 5.20E+01 |
| | **Min** | **8.69E-11** | 6.82E+01 | 2.77E+00 | 1.04E+00 | 1.64E+01 | 4.40E-01 | 1.04E+00 | 2.47E-02 | 6.50E+01 | 6.82E+01 |
| | **Max** | **1.57E-01** | 2.69E+02 | 3.51E+01 | 1.17E+00 | 5.56E+01 | 1.81E+02 | 1.19E+00 | 1.88E+00 | 2.52E+02 | 2.69E+02 |
| **F8** | **Mean** | **1.50E-03** | 4.53E+01 | 5.34E+01 | 2.45E+00 | 9.57E+01 | 4.28E+02 | 6.60E+02 | 1.96E+01 | 6.40E+02 | 4.53E+01 |
| | **Std** | **1.08E-03** | 8.09E+01 | 5.63E+01 | 1.10E+00 | 3.84E+01 | 7.36E+02 | 2.35E+02 | 1.84E+01 | 3.86E+02 | 8.09E+01 |
| | **Min** | **2.82E-04** | 1.87E+00 | 3.90E+00 | 9.35E-01 | 3.62E+01 | 2.51E+00 | 3.02E+02 | 1.41E-01 | 1.93E+02 | 1.87E+00 |
| | **Max** | **5.11E-03** | 4.14E+02 | 2.60E+02 | 6.06E+00 | 1.77E+02 | 3.27E+03 | 1.23E+03 | 6.62E+01 | 1.81E+03 | 4.14E+02 |

**Table 4.17: The Wilcoxon rank sum test results for all benchmark functions with $D$ = 30**

| | **BA** | **DA** | **HS** | **FPA** | **MFO** | **ABC** | **CA** | **GA** | **PSO** |
|---|---|---|---|---|---|---|---|---|---|
| **F1** | 3.02E-11 | 3.02E-11 | 8.99E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 |
| **F2** | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 |
| **F3** | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 |
| **F4** | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 |
| **F5** | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 |
| **F6** | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 |
| **F7** | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 7.38E-10 | 3.02E-11 | 3.02E-11 |
| **F8** | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 |

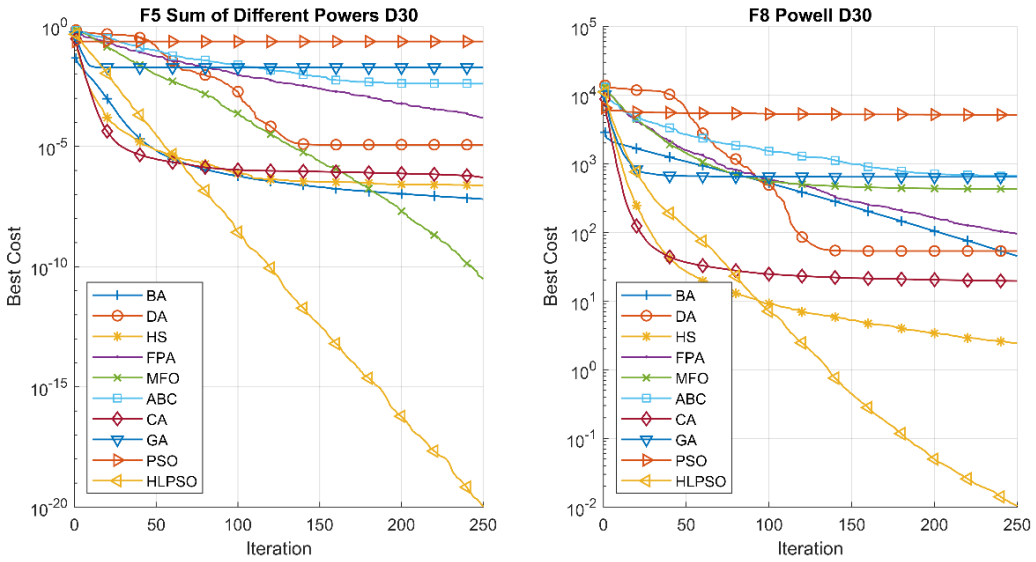



**Figure 4-4: Convergence curves for F5 and F8 in $D$ = 30**

## III. Results for Benchmark Functions with High Dimension ($D$ = 50)

For the high dimensional setting, again the same parameter settings are applied following the experiments for the previous two dimensions 10 and 30 for each method. The outcomes of HLPSO are compared with those of other algorithms. As indicated in Table 4.4, the results of HLPSO are the most effective. Furthermore, for the unimodal and multimodal functions, the increasing of dimension did not significantly impact the HLPSO performance, thus emphasising our optimisation algorithm's scalability. Figure

4-5 presents the convergence curves for various algorithms in this high dimensional setting, while Table 4.19 presents the *p-values* resulting from the Wilcoxon rank sum test. From Table 4.19, it is also evident that HLPSO is a more efficient optimisation algorithm than other compared algorithms for solving high dimensional optimisation problems.

**Table 4.18 : Evaluation results for benchmark functions with $D$ = 50**

| | | HLPSO | BA | DA | HS | FPA | MFO | ABC | CA | GA | PSO |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **F1** | **Mean** | **6.07E+00** | 6.71E+04 | 6.48E+03 | 7.91E+01 | 3.22E+04 | 2.84E+05 | 3.19E+06 | 4.96E+03 | 4.12E+05 | 2.30E+06 |
| | **Std** | **4.09E+00** | 5.39E+04 | 6.75E+03 | 1.87E+01 | 1.27E+04 | 3.90E+05 | 8.10E+05 | 4.51E+03 | 2.49E+05 | 7.71E+05 |
| | **Min** | **7.07E-01** | 2.89E+03 | 9.39E+01 | 4.54E+01 | 1.24E+04 | 8.59E+02 | 1.67E+06 | 5.91E+02 | 1.02E+05 | 1.28E+06 |
| | **Max** | **1.33E+01** | 2.29E+05 | 2.42E+04 | 1.19E+02 | 6.55E+04 | 1.57E+06 | 4.97E+06 | 1.86E+04 | 1.01E+06 | 5.13E+06 |
| **F2** | **Mean** | **5.15E-06** | 1.43E+01 | 9.91E+00 | 3.91E-01 | 1.83E+01 | 1.88E+01 | 1.54E+02 | 1.01E+00 | 4.14E+01 | 1.47E+02 |
| | **Std** | **3.23E-06** | 9.78E+00 | 4.63E+00 | 4.28E-02 | 3.60E+00 | 1.97E+01 | 2.35E+01 | 1.25E+00 | 1.29E+01 | 2.59E+01 |
| | **Min** | **1.40E-06** | 3.54E-01 | 3.73E+00 | 3.09E-01 | 1.19E+01 | 4.86E-01 | 1.06E+02 | 7.09E-02 | 2.31E+01 | 8.46E+01 |
| | **Max** | **1.50E-05** | 3.33E+01 | 2.29E+01 | 4.61E-01 | 2.84E+01 | 7.92E+01 | 1.94E+02 | 6.42E+00 | 8.40E+01 | 1.80E+02 |
| **F3** | **Mean** | **9.20E-04** | 2.75E+05 | 2.16E+04 | 1.68E+03 | 6.83E+04 | 1.11E+05 | 5.05E+05 | 4.85E+03 | 2.40E+05 | 6.03E+05 |
| | **Std** | **7.45E-04** | 8.74E+04 | 1.39E+04 | 4.51E+02 | 1.49E+04 | 9.56E+04 | 7.24E+04 | 4.89E+03 | 4.88E+04 | 1.29E+05 |
| | **Min** | **1.68E-04** | 1.08E+05 | 3.43E+03 | 7.65E+02 | 4.07E+04 | 9.17E+03 | 3.88E+05 | 5.01E+02 | 1.53E+05 | 3.96E+05 |
| | **Max** | **3.06E-03** | 5.05E+05 | 5.79E+04 | 2.73E+03 | 1.01E+05 | 3.83E+05 | 6.54E+05 | 2.33E+04 | 3.87E+05 | 8.79E+05 |
| **F4** | **Mean** | **2.20E-04** | 2.86E+03 | 7.16E+02 | 1.52E+01 | 1.66E+03 | 2.40E+03 | 1.18E+04 | 1.46E+02 | 4.09E+03 | 1.48E+04 |
| | **Std** | **1.79E-04** | 9.49E+02 | 6.90E+02 | 2.16E+00 | 3.90E+02 | 2.35E+03 | 1.62E+03 | 1.91E+02 | 1.25E+03 | 3.97E+03 |
| | **Min** | **4.11E-05** | 1.08E+03 | 2.35E+01 | 1.21E+01 | 9.81E+02 | 1.00E+02 | 8.60E+03 | 5.08E+00 | 2.44E+03 | 7.65E+03 |
| | **Max** | **7.66E-04** | 4.99E+03 | 3.27E+03 | 2.27E+01 | 2.39E+03 | 9.02E+03 | 1.43E+04 | 1.01E+03 | 7.34E+03 | 2.47E+04 |
| **F5** | **Mean** | **3.13E-21** | 5.69E-08 | 6.73E-06 | 4.78E-07 | 2.94E-04 | 2.79E-06 | 7.59E-01 | 6.15E-05 | 2.39E-02 | 3.43E-01 |
| | **Std** | **1.18E-20** | 2.95E-08 | 1.08E-05 | 6.68E-07 | 2.71E-04 | 4.67E-06 | 3.00E-01 | 1.38E-04 | 2.15E-02 | 2.60E-01 |
| | **Min** | **6.46E-26** | 1.13E-08 | 2.52E-17 | 1.08E-08 | 9.78E-06 | 2.74E-08 | 1.87E-01 | 1.86E-07 | 2.92E-03 | 2.36E-02 |
| | **Max** | **5.48E-20** | 1.17E-07 | 4.87E-05 | 3.45E-06 | 9.12E-04 | 1.70E-05 | 1.50E+00 | 6.89E-04 | 8.04E-02 | 9.53E-01 |
| **F6** | **Mean** | **6.21E-01** | 1.57E+01 | 1.01E+01 | 3.09E+01 | 1.38E+01 | 1.85E+01 | 2.03E+01 | 1.10E+01 | 1.61E+01 | 1.95E+01 |
| | **Std** | 7.37E-01 | 9.40E-01 | 2.03E+00 | 3.97E-01 | 9.09E-01 | 1.60E+00 | **2.54E-01** | 2.36E+00 | 8.56E-01 | 5.71E-01 |
| | **Min** | **8.06E-04** | 1.37E+01 | 5.58E+00 | 3.02E+01 | 1.18E+01 | 1.31E+01 | 1.98E+01 | 5.46E+00 | 1.44E+01 | 1.82E+01 |
| | **Max** | **1.81E+00** | 1.74E+01 | 1.38E+01 | 4.08E+01 | 1.53E+01 | 1.97E+01 | 2.08E+01 | 1.74E+01 | 1.81E+01 | 2.05E+01 |
| **F7** | **Mean** | **1.13E-02** | 2.31E+02 | 2.95E+01 | 3.06E+00 | 6.38E+01 | 6.58E+01 | 5.59E+02 | 5.77E+00 | 2.56E+02 | 5.28E+02 |
| | **Std** | **1.44E-02** | 6.55E+01 | 1.73E+01 | 4.47E-01 | 1.30E+01 | 6.33E+01 | 7.47E+01 | 5.48E+00 | 7.11E+01 | 9.93E+01 |
| | **Min** | **2.22E-06** | 1.06E+02 | 6.04E+00 | 2.08E+00 | 3.96E+01 | 3.51E+00 | 3.73E+02 | 1.15E+00 | 1.45E+02 | 3.09E+02 |
| | **Max** | **4.69E-02** | 4.14E+02 | 7.76E+01 | 3.78E+00 | 9.35E+01 | 1.87E+02 | 7.23E+02 | 2.83E+01 | 4.39E+02 | 6.91E+02 |
| **F8** | **Mean** | **9.28E-02** | 2.43E+02 | 3.58E+02 | 2.25E+01 | 3.98E+02 | 2.98E+03 | 1.56E+04 | 1.49E+02 | 2.35E+03 | 1.40E+04 |
| | **Std** | **8.20E-02** | 3.06E+02 | 5.01E+02 | 8.61E+00 | 9.86E+01 | 2.01E+03 | 3.57E+03 | 1.14E+02 | 1.10E+03 | 4.03E+03 |
| | **Min** | **2.98E-02** | 1.67E+01 | 3.04E+01 | 8.69E+00 | 2.04E+02 | 1.76E+02 | 1.10E+04 | 1.99E+01 | 1.02E+03 | 6.44E+03 |
| | **Max** | **3.71E-01** | 1.57E+03 | 2.49E+03 | 4.77E+01 | 6.25E+02 | 7.90E+03 | 2.35E+04 | 4.62E+02 | 6.55E+03 | 2.24E+04 |

**Table 4.19: The Wilcoxon rank sum test results for all benchmark functions with $D$ = 50**

| | BA | DA | HS | FPA | MFO | ABC | CA | GA | PSO |
|---|---|---|---|---|---|---|---|---|---|
| **F1** | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 |
| **F2** | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 |
| **F3** | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 |
| **F4** | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 |
| **F5** | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 |
| **F6** | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 |
| **F7** | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 |
| **F8** | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 |

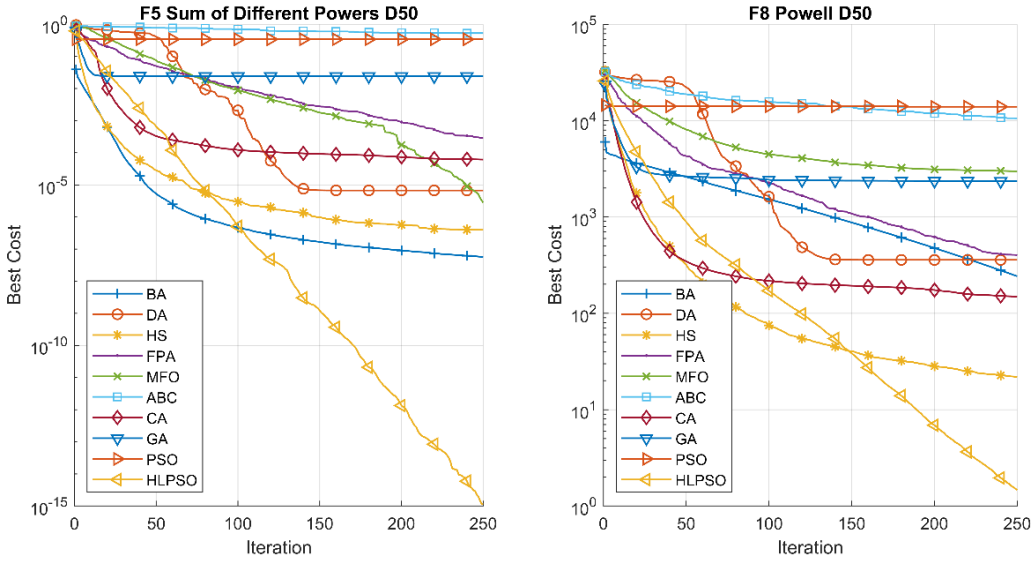



**Figure 4-5: Convergence curves for F5 and F8 in $D$ = 50**

## 4.4 Chapter Summary

This chapter presents an enhanced decision support system with the aim of identifying benign and malignant skin lesions in dermoscopy images by proposing a new PSO model, i.e. HLPSO. One of the key benefits associated with such an initiative intelligent system is the way in which it represents significant potential in regarding the opportunity to enhance the precision of skin lesion classification. Recently, PSO has been considered as a novel and favourable optimisation algorithm because it is easy to use and efficient, but conventional PSO has several limitations. Notable challenges arise from the issues of premature convergence and becoming stuck in local optima. Therefore, these problems could affect the model's search capabilities for complex optimization problems. This chapter has addressed this limitation by proposing and evaluating a novel PSO variant that can be employed for the optimisation of skin lesion identification. It incorporates a series of search methods that have been effectively employed in the context of numerous medical imaging areas.

For the evaluation of several skin lesion dataset, the outcomes reveal that HLPSO achieves superior results, particularly in terms of convergence speed, global optimality and the precision of results. Therefore, continued research in this area should focus on the

advancement and perfection of a novel hybrid feature selection algorithm, and this could be accomplished by integrating the PSO with alternative search algorithms such as FA with adaptive search parameters. The principal advantage of this is that it would facilitate the identification of more discriminative features, thereby enhancing the extent to which classification is accurate.

# CHAPTER 5
# PROPOSED METHODOLOGY 3 WITH PSOVA BASED FEATURE SELECTION

To achieve the best trade-off between swarm diversity and convergence speed, this chapter puts forward another PSO variant, namely the PSOVA model, which employs food attraction and enemy evading actions, the sub-swarm concept, mutation strategies and dynamic matrix representations to determine the significant features for skin lesion classification. In comparison with other algorithms, PSOVA includes additional enemy avoidance strategies, which efficiently avoid worse experiences while moving towards the optimal regions. This algorithm is able to explore a wider search space, therefore, has better chances of finding the global optima, in comparison with other search algorithms. Furthermore, PSOVA stands out from other cutting-edge PSO-based optimisation methods by enhancing the exploration of both the swarm leader and the population; this is to capture the global optimum using a cooperative strategy through collaboration of the sub-swarm and the mutation operations with probability distributions. Addressing the premature convergence issue of conventional PSO, this algorithm is capable of effectively choosing features discriminatively and employs the actions of survival tactics to increase convergence speed.

## 5.1 The Proposed PSOVA Search Strategies for Feature Selection

The proposed PSOVA algorithm updated the velocity using both the original PSO attraction operation and an additional flee action. The PSOVA algorithm first initialised the search space with a swarm of 50 particles. An identical series of iteration steps are repeatably conducted to identify the optimal feature subset, and this is carried out to the satisfaction of the maximum number of iterations. In each iteration, each particle determines the position of personal best, $p_{best}$, and the global best $g_{best}$ of the overall swarm using both attraction and evading actions. A KNN-based classifier is employed to assess particles and provides these particles with fitness values. It is essential to ensure the particle's position never exceeds either the predefined lowest or highest boundary parameters. The selection of the best leader and three worst particles from the population will be conducted, and the minimal correlation to the $g_{best}$ will be used to search for the second leader. If the second leader has been identified by using the correlation Equation (5.1), the swarm will be separated into two sub-swarms, i.e. ss1 and ss2, from the initial swarm.

One leader will lead each sub-swarm, thus expanding the search space and enhancing both local and global search capacity. In each sub-swarm, if a particle is one of those three worst particles previously identified, the conventional PSO – as shown in Equation (3.7) – will be used to update the particle's velocity in each dimension. The remaining particles in the group will perform the following four search methods for leading to the optimum solution, i.e. 1) each particle follows the leader fully in every dimension as defined in the original PSO algorithm, 2) it randomly selects some sub-dimensions and follows the leader partially in those dimensions, 3) it avoids personal and global worst experiences in every dimension, 4) it randomly selects some sub-dimensions and avoids personal and global worst individuals partially in those dimensions.

These four search mechanisms increase diversification of the search process and enable the proposed PSO model to explore distinctive search regions and to reduce the likelihood of being trapped in local optima. During the search process, each particle performs each

of the four mechanisms for position updating. Then, the best offspring stems from the four actions is employed for velocity updating. After the subswarm-based search process iterates for a number of iterations, a new subswarm leader is retrieved in each subpopulation. The new subswarm leader is further enhanced using three probability distributions, i.e. Gaussian, Cauchy, and Levy distributions from Equation (4.2)-(4.4) are used. If the offspring solution generated by any of these random walk strategies has a better fitness score than that of the current parent subswarm leader, this promising offspring is used to replace the parent subswarm leader. Finally, the two subpopulations are merged, i.e. $ss1$ and $ss2$, into one population which contains a slightly different set of particles compared with the original population, and the best solution among the two subswarm leaders is regarded as the new global best solution. Another set of three worst solutions is also retrieved and used to update the worst archive.

However, if the second swarm leader, which possesses a promising fitness score but embeds a low correlation to the global best solution, cannot be retrieved at the beginning stage of the search process in a certain iteration, the search operations described above are conducted purely using the primary swarm, instead of two subswarms.

Subsequently, the matrix representation of the swarm is dynamically adjusted to a new form by switching the rows and columns of the original matrix (e.g. changing a particle matrix representation from 5×29 to 29×5) and the overall search process is repeated using this new matrix representation. The algorithm iterates until the termination criteria are reached. The pseudo-code for the PSOVA algorithm is depicted in Algorithm 4.

**Algorithm 5.1: Pseudo-code of The Proposed PSOVA model**

| | |
|---|---|
| **1** | **Start** |
| **2** | Initialize particles (e.g. 50 particles); |
| **3** | Evaluate the population to identify the initial best leader $gBest$; |
| **4** | Copy the best leader $gBest$ to the **best memory**; |
| **5** | **While** (! Stop condition) do {// 5 iterations |
| **6** | **For** (each particle $x_i$ in the particle) **do** { |
| **7** | **If** (\|fitness_ $x_i$ – fitness_ $gBest$ \| < Threshold Best) && (correlation between $x_i$and $gBest$ < 0) {//Similar in fitness values but more than 50% different in positions |
| **8** | Select the candidate, $x_i$, as the second leader; |
| **9** | } **End If** |

| | |
|---|---|
| **10** | } **End For** |
| **11** | Identify three worst solutions in the population and store them in the **worst memory**; |
| **12** | **If** (Two leaders) { |
| **13** | Randomly separate the population into two groups with each group led by one leader; // Sub-swarm 1 led by $gBest$ and sub-swarm two led by the second best leader |
| **14** | **While** (! Stop condition) do {// 24 iterations |
| **15** | **For** (each group) **do** { |
| **16** | **For** (each particle $x_i$ in the group) **do** { |
| **17** | **If** ($x_i$ is one of the worst solutions in the **worst memory**) { |
| **18** | Follow the leader fully with Eq. (3.7) & (3.8) in every dimension; |
| **19** | } **Else If** {// xi ≠ leader or the worst solutions |
| **20** | Run all steps below and decide which has the best solution; |
| **21** | 1. Partially avoid the worst solutions with Eq. (5.2)-(5.4) (by randomly identifying some sub-dimensions in a row/column/diagonal in the matrix) and flying away from the worst in those dimensions; |
| **22** | 2. Partially follow towards the leader with Eq. (3.7) & (3.8) (by randomly identifying some sub-dimensions and follow the leader in those sub-dimensions); |
| **23** | 3. Fully avoid the worst solutions with Eq. (5.2)-(5.4) in every dimension; |
| **24** | 4. Fully follow the leader with Eq. (3.7) & (3.8) in every dimension; |
| **25** | } **End If** |
| **26** | Evaluate xi and update $pBest$; |
| **27** | } **End For** |
| **28** | Evaluate the particles in the sub-swarm and update the sub-swarm leader; |
| **29** | Conduct long jumps of the current swarm leader using Levy flights/Gaussian/Cauchy distributions defined in Eq. (4.3)-(4.5); |
| **30** | Update the current swarm leader if the new position developed by long jumps has better fitness; |
| **31** | } **End For** |
| **32** | Compare two sub-swarm leaders and store the best sub-swarm leader; |
| **33** | } **End While**; |
| **34** | Combine the groups and update the best leader, $gBest$, in the **best memory** with the best sub-swarm leader; |
| **35** | } **Else** (There is just one leader) { |
| **36** | Conduct the single group optimisation using lines 14-34; |
| **37** | Update the best leader, $gBest$, in the **best memory**; |
| **38** | } **End If** |
| **39** | } **End While** |
| **40** | Change the matrix representation; |
| **41** | **While** (! Stop condition) do {// 5 iterations |
| **42** | Repeat line 5-40; |
| **43** | } **End While** |
| **44** | Return the most optimal solution; |
| **45** | **End** |

### 5.1.1 The Selection of the Second Leader

The correlation relationships between the leader and other particles have influenced the choice of the second swarm leader, as shown in Algorithm 4 line 7. These relationships are based on both the distance and the fitness variations between two particles A (the swarm leader) and B. The correlation $corr2$ function from MATLAB has been utilised to find the correlation relationship between a pair of particles, and this was established by Equation (5.1) (Snedecor and Cochran, 1967). The function also uses the histogram correlation comparison technique for the detection of particles most or least correlated with the leader.

$$corr(A,B) = \frac{\sum_m \sum_n (A_{mn} - \bar{A})(B_{mn} - \bar{B})}{\sqrt{(\sum_m \sum_n (A_{mn} - \bar{A})^2)(\sum_m \sum_n (B_{mn} - \bar{B})^2)}} \quad (5.1)$$

The correlation relationship of two particles is denoted by $corr$, the swarm leader being denoted by A and the follower particle being represented by B. The output of Equation (5.1) is in the range -1 to 1, with the minimal correlation denoted by -1 and the maximal correlation denoted by 1. After identifying the global best solution and the second swarm leader, it divide the overall swarm into two subswarms. The search process of each subswarm is guided by each leader. Since the two leaders are remote in positions, i.e. a low correlation in position, it is more likely that they lead the subswarm-based search process to explore distinctive regions, in an attempt to avoid stagnation. For example, if the minimal correlation of -1 is found, the second swarm leader is identified with similar fitness but remote in position, and the population will be divided into 2 sub-swarms. If the maximal correlation of 1 is found, the population will not be divided owing to the lack of a competitive particle with sufficient difference in position to the leader.

### 5.1.2 Attraction and Avoidance Actions

The particles followed by using conventional PSO operation tend to be trapped in local optima since the search is led by a single leader. Therefore, the proposed algorithm utilises both attraction and an additional avoidance search technique, together with sub-swarm-based mechanism to increase local or global exploration and search diversity. Instead of purely following the sub-swarm leader in each dimension, the PSOVA

algorithms also enables the search to follow the sub-swarm leader in any randomly selected sub-dimensions, e.g. a randomly selected row, column or diagonal. In each iteration, the proposed algorithm detects and stores several worst solutions in the worst memory; this includes the local worst, *pWorst* and three global worst solutions, i.e. the *gWorst* solutions. They are used as part of the avoidance search mechanism. The following Equation (5.2)-(5.4) has been used to avoid the worst experiences, in each dimension or any randomly selected sub-dimensions, which is identical to the above attraction action. $p_{iw}$ designates the *pWorst* solution and $p_{gw}$ indicates the average solution from the three *gWorst* solutions, which are denoted as *w1, w2* and *w3*. We have also used the same parameter settings, e.g. for the initial weight and acceleration coefficients, *C1* and *C2*, for this avoidance mechanism. The detailed parameter setting of the PSOVA model can be found in Section 5.2.1. The model uses the following defined evading actions to lead the search to avoid unpromising regions.

$$v_{id}^{t+1} = w * v_{id}^{t} - c1 * r1 * (p_{iw} - x_{id}^{t}) - c2 * r2 * (p_{gw} - x_{id}^{t}) \quad (5.2)$$

$$p_{iw} = \frac{w1 + w2 + w3}{3} \quad (5.3)$$

$$x_{id}^{t+1} = x_{id}^{t} + v_{id}^{t+1} \quad (5.4)$$

In the original PSO model, the search process is purely guided by the swarm leader. When the attraction driven search mechanism stagnates, this results in premature convergence of the original PSO model. On the contrary, the proposed evading action is complementary to the original attraction operation. In the proposed PSO algorithm, when the attraction mechanism guided by the subswarm leader becomes stagnant, the evading operation pushes the subswarm particles away from less optimal regions to overcome stagnation.

### 5.1.3 Dynamic Matrix Representations

Finally, the matrix representation is dynamically adjusted to increase the search diversity; this involves randomly adjusting the column and row of the swarm particle representation, i.e. changing the particle representation using 50 columns × 10 rows, to

e.g. 10 columns × 50 rows, 125 columns × 4 rows or 4 columns × 125 rows. The search processes discussed in Sections 0-5.1.2 are conducted again under the new matrix representation. In addition, during the search process, when the algorithm cannot identify a second swarm leader that has a similar fitness score but with the least in correlation to the swarm leader, the search processes presented in Section 5.1.2 are conducted using a primary swarm led by a single leader, instead of two subswarm leaders.

## 5.2 Evaluation

Two data sets from Chapter 4, i.e. PH2 and Dermofit, are also used in this study for evaluating the newly proposed PSOVA model, Also, an additional third dataset from Dermnet (2017) has been added for evaluation. The Dermnet dataset has a total of 152 images with 45 benign and 107 malignant melanomas. In this study, we used all the malignant melanoma images from the Dermnet dataset for experimentation. Therefore, in total, the combined dataset using the above three databases contains 207 benign images with 80 from PH2, and 127 from Dermofit, and 223 malignant images with 40 from PH2, 76 from Dermofit and 107 from Dermnet. The reason of this combined dataset is selected is due to the imbalance class problems in existing datasets. This combined dataset contains more balance samples for each class and has been used to train the proposed model. In detail, this research utilised an 80:20 aspect ratio for the separation of training and testing, which included 344 images, (166 benign and 178 melanomas), and 86 images (41 benign and 45 melanomas) for training and testing, respectively. Specifically, the training set comprises 64 PH2 and 102 Dermofit benign images; and 32 PH2, 61 Dermofit and 85 Dermnet melanoma images. Moreover, the testing set consists of 16 PH2 and 25 Dermofit benign images, and 8 PH2, 15 Dermofit and 22 Dermnet melanoma images.

**Table 5.1: Data sets for skin cancer classification**

| Dataset | Benign | Melanoma |
|---|---|---|
| **PH2** | 80 | 40 |
| **Dermofit** | 127 | 76 |
| **Dermnet** | N/A | 107 |

For the experiment, it evaluated the performance of classification with two classifiers including KNN and SVM combined with both 10-fold and hold-out validations. In this

research, all the testing was conducted with 30 runs for each method to ensure the fairness of the comparison. Since each method has different internal search strategies, a maximum of 50 population × 500 iterations = 25,000 function evaluations are used as the stopping criteria. Despite the existence of multiple search strategies inherent to the PSOVA model, all utilising an identical number of function evaluations over the training phase, i.e. 53 particles× 120 iterations × 4 function evaluations = 25,440 function evaluations, for performance comparison. Overall, the experiment employs an identical number of function evaluations (i.e. 50 particles × 500 iterations) used for all classical search methods and other elite PSO variants such as DNLPSO and BBPSOV, to enable a fair comparison.

In this section, the baseline algorithms previously used for comparison in Section 4.3 are also employed here. Several additional modified PSO methods such as BBPSO (Krohling and Mendel, 2009), DNLPSO (Nasir et al., 2012), ELPSO (Jordehi, 2015), AGPSO (Mirjalili et al., 2014a), ThBPSO (Krisshna et al., 2014), MS-PSO (Yang et al., 2017), BBPSOV (Srisukkham et al., 2017) and GPSO (Chen et al., 2016) have been added to enhance the comparison with the proposed model.

### 5.2.1 Parameter Settings

For the verification purpose of the algorithms and the experimental results, the proposed PSOVA algorithm is compared against the conventional algorithms and advanced PSO variants for benchmarking the performance. Based on experimental trials, the modified PSO operation is associated with the following settings. The employed parameter of inertia weight is 0.99 to identify the effect of the prior speed on repetitive searches. As with the initial PSO algorithm, acceleration constants ($C1 = C2 = 2.5$) are employed to expedite convergence. For other search methods, the parameter settings are employed as defined in their original studies.

**Table 5.2: Parameter settings of the proposed and other methods**

| Algorithms | Parameters |
|---|---|
| **BBPSO** | No parameter setting required |
| **DNLPSO** | *C1*=*C2*=1.49445, refreshing gap=3, regrouping period=5, inertia weight=0.9-(0.9-0.4) x (k-1)/(max_gen-1), where k and max_gen represent the current and maximum iteration number, respectively. |
| **ELPSO** | Maximum velocity=0.6, inertia weight=0.9, acceleration constants *C1*=*C2*=2 |
| **AGPSO** | Maximum velocity=0.6, inertia weight=0.9, adaptive decreasing *C1* and increasing *C2* over generations |
| **ThBPSO** | Maximum velocity=0.6, inertia weight=0.9, acceleration constants *C1*=*C2*=2 |
| **MS-PSO** | Maximum velocity=0.6, inertia weight=0.9, acceleration constants *C1*=*C2*=2 |
| **GM-PSO** | Maximum velocity=0.6, inertia weight=0.9, acceleration constants *C1*=*C2*=2, crossover probability=0.7, mutation probability=0.3 |
| **BBPSOV** | The logistic map used as the search parameter |
| **GPSO** | Maximum velocity=0.6, inertia weight=0.9, acceleration constants *C1*=*C2*=2, crossover probability=0.7, mutation probability=0.3 |
| **PSOVA** | Maximum velocity=0.6, inertia weight=0.99, acceleration constants *C1*=*C2*=2.5 |

### 5.2.2 Evaluation Using the Combined Dataset

We used a combined dataset for the evaluation with a total of 30 runs for each method. The multiple runs were to find out the best feature subsets from each algorithm. As seen in Table 5.4, in comparison with all other algorithms, the proposed method achieved the highest average geometric mean rates when combined with KNN and SVM classifiers for both 10-fold and hold-out validations. Based on the classification results with KNN using 10-fold validation, PSOVA has outperformed BA, DA, HS, FPA, MFO, ABC, CA, GA, PSO, BPSO, ThBPSO, BPSOV, AGPSO, ELPSO, DNLPSO, GM-PSO, MS-PSO and GPSO, respectively. Furthermore, for KNN with the hold-out validation, this experiment outperformed BA, DA, HS, FPA, MFO, ABC, CA, GA, PSO, BPSO, ThBPSO, BPSOV, AGPSO, ELPSO, DNLPSO, GM-PSO, MS-PSO and GPSO, respectively. When the SVM-based classifier was applied with 10-fold validation, the proposed algorithm outperformed BA, DA, HS, FPA, MFO, ABC, CA, GA, PSO, BPSO, ThBPSO, BPSOV, AGPSO, ELPSO, DNLPSO, GM-PSO, MS-PSO and GPSO, respectively. And lastly, with the SVM and the hold-out classification, PSOVA outperforms BA, DA, HS, FPA, MFO, ABC, CA, GA, PSO, BPSO, ThBPSO, BPSOV, AGPSO, ELPSO, DNLPSO, GM-PSO, MS-PSO and GPSO, respectively. The Wilcoxon rank sum test results shown in Table 5.4 also further indicate the superiority of the proposed PSOVA model over other search methods.

**Table 5.3: Average performance correlation of each algorithm over 30 runs for the combined dataset**

| | KNN | | SVM | |
|---|---|---|---|---|
| | **10-fold** | **Hold-out** | **10-fold** | **Hold-out** |
| **BA** | 2.22% | 1.87% | 1.91% | 1.67% |
| **DA** | 2.39% | 1.90% | 1.27% | 1.03% |
| **HS** | 0.96% | 0.89% | 0.89% | 0.87% |
| **FPA** | 2.64% | 2.24% | 1.53% | 1.37% |
| **MFO** | 2.63% | 2.20% | 3.76% | 3.49% |
| **ABC** | 2.51% | 2.27% | 2.15% | 2.07% |
| **CA** | 1.94% | 1.77% | 1.55% | 1.44% |
| **GA** | 4.23% | 3.88% | 6.05% | 5.76% |
| **PSO** | 2.73% | 2.41% | 1.88% | 1.70% |
| **BBPSO** | 1.96% | 1.63% | 2.15% | 1.96% |
| **ThBPSO** | 3.85% | 3.45% | 6.18% | 5.70% |
| **BBPSOV** | 2.09% | 1.69% | 3.09% | 2.88% |
| **AGPSO** | 2.22% | 1.94% | 2.07% | 2.01% |
| **ELPSO** | 1.70% | 1.53% | 2.58% | 2.46% |
| **DNLPSO** | 2.46% | 2.05% | 2.15% | 1.98% |
| **GM-PSO** | 2.73% | 2.32% | 2.63% | 2.51% |
| **MS-PSO** | 2.72% | 2.45% | 2.84% | 2.61% |
| **GPSO** | 1.77% | 1.51% | 2.54% | 2.27% |

**Table 5.4: Average performance of each algorithm over 30 runs for the combined dataset**

| | KNN | | SVM | |
|---|---|---|---|---|
| | **10-fold** | **Hold-out** | **10-fold** | **Hold-out** |
| **BA** | 0.9009 | 0.9075 | 0.9232 | 0.9284 |
| **DA** | 0.8992 | 0.9072 | 0.9296 | 0.9348 |
| **HS** | 0.9135 | 0.9173 | 0.9334 | 0.9364 |
| **FPA** | 0.8967 | 0.9038 | 0.9270 | 0.9314 |
| **MFO** | 0.8968 | 0.9042 | 0.9047 | 0.9102 |
| **ABC** | 0.8980 | 0.9035 | 0.9208 | 0.9244 |
| **CA** | 0.9037 | 0.9085 | 0.9268 | 0.9307 |
| **GA** | 0.8808 | 0.8874 | 0.8818 | 0.8875 |
| **PSO** | 0.8958 | 0.9021 | 0.9235 | 0.9281 |
| **BBPSO** | 0.9035 | 0.9099 | 0.9208 | 0.9255 |
| **ThBPSO** | 0.8846 | 0.8917 | 0.8805 | 0.8881 |
| **BBPSOV** | 0.9022 | 0.9093 | 0.9114 | 0.9163 |
| **AGPSO** | 0.9009 | 0.9068 | 0.9216 | 0.9250 |
| **ELPSO** | 0.9061 | 0.9109 | 0.9165 | 0.9205 |
| **DNLPSO** | 0.8985 | 0.9057 | 0.9208 | 0.9253 |
| **GM-PSO** | 0.8958 | 0.9030 | 0.9160 | 0.9200 |
| **MS-PSO** | 0.8959 | 0.9017 | 0.9139 | 0.9190 |
| **GPSO** | 0.9054 | 0.9111 | 0.9169 | 0.9224 |
| **PSOVA** | **0.9231** | **0.9262** | **0.9423** | **0.9451** |

**Table 5.5: The *p-values* of the Wilcoxon rank sum test using the combined dataset**

| | **BA** | **DA** | **HS** | **FPA** | **MFO** | **ABC** | **CA** | **GA** | **PSO** |
|---|---|---|---|---|---|---|---|---|---|
| **KNN Wilcoxon 10F** | 1.73E-04 | 5.26E-05 | 2.51E-05 | 1.13E-05 | 3.37E-04 | 6.55E-05 | 9.26E-03 | 3.96E-08 | 6.73E-06 |
| **KNN Wilcoxon HO** | 2.44E-04 | 1.79E-04 | 3.00E-05 | 2.93E-05 | 5.08E-04 | 1.96E-05 | 3.09E-03 | 2.88E-08 | 1.12E-05 |
| **SVM Wilcoxon 10F** | 3.11E-03 | **6.79E-02** | 6.91E-04 | 5.20E-03 | 1.47E-07 | 3.16E-05 | 7.78E-03 | 8.99E-11 | 4.43E-03 |
| **SVM Wilcoxon HO** | 1.38E-03 | 4.69E-02 | 2.23E-03 | 1.10E-02 | 2.86E-07 | 2.92E-05 | 5.95E-03 | 1.87E-10 | 5.66E-03 |

| | BBPSO | ThBPSO | BBPSOV | AGPSO | ELPSO | DNLPSO | GM-PSO | MSPSO | GPSO |
|---|---|---|---|---|---|---|---|---|---|
| **KNN Wilcoxon 10F** | 3.56E-04 | 8.35E-08 | 2.28E-05 | 2.44E-03 | 3.76E-04 | 2.19E-04 | 1.79E-06 | 7.04E-07 | 1.63E-04 |
| **KNN Wilcoxon HO** | 6.51E-04 | 9.84E-08 | 7.80E-05 | 2.69E-03 | 2.88E-04 | 2.95E-04 | 1.84E-06 | 8.96E-07 | 2.29E-04 |
| **SVM Wilcoxon 10F** | 1.17E-04 | 2.02E-08 | 1.58E-04 | 1.64E-03 | 1.41E-04 | 2.44E-03 | 3.09E-04 | 1.73E-04 | 1.95E-04 |
| **SVM Wilcoxon HO** | 4.90E-04 | 3.50E-08 | 7.60E-05 | 1.30E-03 | 1.85E-04 | 2.21E-03 | 7.83E-04 | 1.37E-04 | 1.84E-04 |

**Table 5.6: The training computational cost and the average number of features selected using the combined dataset**

| | PSOVA | BA | DA | HS | FPA | MFO | ABC | CA | GA | PSO |
|---|---|---|---|---|---|---|---|---|---|---|
| **Average no. of selected features** | 52.11 | 74.35 | 78.09 | 71.86 | 70.73 | 65.93 | 70.78 | 67.38 | 62.69 | 73.58 |
| **Training cost (seconds)** | 3908.46 | 3816.57 | 3813.44 | 3931.89 | 3806.12 | 3905.52 | 3906.70 | 3929.63 | 3791.70 | 3878.16 |

| | PSOVA | BBPSO | ThBPSO | BBPSOV | AGPSO | ELPSO | DNLPSO | GM-PSO | MSPSO | GPSO |
|---|---|---|---|---|---|---|---|---|---|---|
| **Average no. of selected features** | 52.11 | 73.17 | 53.13 | 64.67 | 65.31 | 77.33 | 69.73 | 74.26 | 59.62 | 72.37 |
| **Training cost (seconds)** | 3908.46 | 3937.50 | 4048.19 | 3981.04 | 3879.13 | 3970.20 | 3860.92 | 3878.48 | 3998.31 | 3945.48 |

### 5.2.3 Evaluation Using the PH2 Dataset

The assessment was also performed on an experimental group of 200 images from PH2 dataset. The average classification performance of 30 runs relevant to every optimisation algorithm in conjunction with various classifiers is summarised in Table 5.8. The best results were correlated with the use of SVM in combination with each feature selection model. The experiment using KNN has demonstrated that the mean GM performances of PSOVA for two lesion classes are 94.16% for 10-fold and, 96.04% for hold-out respectively. These tests gave more effective outcomes when contrasted with other algorithms, and our model outperforms BA, DA, FPA, HS, MFO, CA, ABC, GA, PSO, BBPSO, ThBPSO, ELPSO, BBPSOV, AGPSO, DNLPSO, GM-PSO, MS-PSO and GPSO respectively for 10-fold validation, while for hold-out validation, it outperformed BA, DA, FPA, HS, MFO, CA, ABC, GA, PSO, BBPSO, ThBPSO, ELPSO, BBPSOV, AGPSO, DNLPSO, GM-PSO, MS-PSO and GPSO, respectively. Moreover, in SVM, the two-class lesion classification using PSOVA achieved 95.23% with 10-fold validation and 96.45% for hold-out validation, which were higher than the results obtained using the KNN classification. It also outperformed the BA, DA, FPA, HS, MFO, CA, ABC, GA, PSO, BBPSO, ThBPSO, ELPSO, BBPSOV, AGPSO, DNLPSO, GM-PSO, MS-PSO and GPSO respectively for 10-fold validation; and for hold-out validation, it outperformed the BA, DA, FPA, HS, MFO, CA, ABC, GA, PSO, BBPSO, ThBPSO, ELPSO, BBPSOV, AGPSO, DNLPSO, GM-PSO, MS-PSO and GPSO, respectively.

**Table 5.7: Average performance correlation of each algorithm over 30 runs for the PH2 dataset**

| | KNN | | SVM | |
|---|---|---|---|---|
| | **10-fold** | **Hold-out** | **10-fold** | **Hold-out** |
| **BA** | 4.71% | 4.17% | 4.12% | 1.89% |
| **DA** | 5.71% | 3.95% | 3.57% | 2.28% |
| **HS** | 6.63% | 6.75% | 1.79% | 1.05% |
| **FPA** | 5.46% | 3.57% | 0.91% | 0.31% |
| **MFO** | 0.74% | 0.80% | 2.07% | 0.83% |
| **ABC** | 3.21% | 3.14% | 2.85% | 0.71% |
| **CA** | 3.66% | 3.48% | 1.83% | 0.62% |
| **GA** | 2.16% | 2.40% | 3.02% | 1.60% |
| **PSO** | 5.73% | 3.87% | 4.30% | 1.77% |
| **BBPSO** | 5.53% | 4.43% | 4.29% | 2.71% |
| **ThBPSO** | 6.15% | 7.35% | 5.19% | 5.91% |
| **BBPSOV** | 1.61% | 1.79% | 2.79% | 1.26% |
| **AGPSO** | 3.27% | 2.43% | 3.04% | 2.17% |
| **ELPSO** | 2.38% | 2.56% | 2.57% | 1.60% |
| **DNLPSO** | 3.59% | 2.77% | 4.14% | 2.67% |
| **GM-PSO** | 3.69% | 3.32% | 3.37% | 2.31% |
| **MS-PSO** | 4.29% | 2.48% | 3.27% | 2.11% |
| **GPSO** | 2.93% | 2.31% | 5.22% | 2.58% |

**Table 5.8: Average performance of each algorithm over 30 runs for the PH2 dataset**

| | KNN | | SVM | |
|---|---|---|---|---|
| | **10-fold** | **Hold-out** | **10-fold** | **Hold-out** |
| **BA** | 0.8945 | 0.9187 | 0.9111 | 0.9456 |
| **DA** | 0.8845 | 0.9209 | 0.9166 | 0.9417 |
| **HS** | 0.8753 | 0.8929 | 0.9344 | 0.9540 |
| **FPA** | 0.8870 | 0.9247 | 0.9432 | 0.9614 |
| **MFO** | 0.9342 | 0.9524 | 0.9316 | 0.9562 |
| **ABC** | 0.9095 | 0.9290 | 0.9238 | 0.9574 |
| **CA** | 0.9050 | 0.9256 | 0.9340 | 0.9583 |
| **GA** | 0.9200 | 0.9364 | 0.9221 | 0.9485 |
| **PSO** | 0.8843 | 0.9217 | 0.9093 | 0.9468 |
| **BBPSO** | 0.8863 | 0.9161 | 0.9094 | 0.9374 |
| **ThBPSO** | 0.8801 | 0.8869 | 0.9004 | 0.9054 |
| **BBPSOV** | 0.9255 | 0.9425 | 0.9244 | 0.9519 |
| **AGPSO** | 0.9089 | 0.9361 | 0.9219 | 0.9428 |
| **ELPSO** | 0.9178 | 0.9348 | 0.9266 | 0.9485 |
| **DNLPSO** | 0.9057 | 0.9327 | 0.9109 | 0.9378 |
| **GM-PSO** | 0.9047 | 0.9272 | 0.9186 | 0.9414 |
| **MS-PSO** | 0.8987 | 0.9356 | 0.9196 | 0.9434 |
| **GPSO** | 0.9123 | 0.9373 | 0.9001 | 0.9387 |
| **PSOVA** | **0.9416** | **0.9604** | **0.9523** | **0.9645** |

**Table 5.9: The *p-values* of the Wilcoxon rank sum test using PH2 dataset**

| | **BA** | **DA** | **HS** | **FPA** | **MFO** | **ABC** | **CA** | **GA** | **PSO** |
|---|---|---|---|---|---|---|---|---|---|
| **KNN Wilcoxon 10F** | 5.10E-03 | 1.60E-03 | 9.76E-03 | 4.82E-03 | 2.81E-04 | 1.07E-04 | 6.15E-05 | 1.54E-04 | 2.06E-03 |
| **KNN Wilcoxon HO** | 8.85E-04 | 1.91E-03 | 7.38E-03 | 3.45E-03 | 2.07E-04 | 3.43E-04 | 5.42E-05 | 2.69E-05 | 1.60E-03 |
| **SVM Wilcoxon 10F** | 4.90E-03 | 5.37E-03 | 2.08E-03 | 3.80E-04 | 2.52E-04 | 1.28E-04 | 5.20E-03 | 5.69E-05 | 9.58E-03 |
| **SVM Wilcoxon HO** | 1.00E-02 | 8.45E-03 | 4.93E-03 | 8.41E-04 | 2.39E-03 | 1.61E-04 | 4.29E-03 | 2.73E-04 | 9.20E-03 |

| | BBPSO | ThBPSO | BBPSOV | AGPSO | ELPSO | DNLPSO | GM-PSO | MSPSO | GPSO |
|---|---|---|---|---|---|---|---|---|---|
| **KNN Wilcoxon 10F** | 3.21E-03 | 2.69E-05 | 1.73E-03 | 5.89E-03 | 4.11E-03 | 9.70E-03 | 6.44E-03 | 9.53E-03 | 8.53E-03 |
| **KNN Wilcoxon HO** | 7.28E-04 | 1.19E-06 | 1.98E-03 | 4.99E-03 | 5.67E-03 | 1.00E-02 | 7.53E-04 | 6.13E-03 | 6.56E-03 |
| **SVM Wilcoxon 10F** | 8.74E-03 | 1.08E-05 | 2.94E-03 | 4.01E-03 | 1.94E-03 | 7.25E-03 | 8.60E-05 | 3.90E-03 | 9.28E-03 |
| **SVM Wilcoxon HO** | 3.86E-03 | 1.58E-05 | 4.79E-03 | 9.06E-03 | 2.61E-03 | 4.48E-03 | 1.04E-04 | 9.56E-03 | 7.28E-03 |

**Table 5.10: The training computational cost and the average number of features selected using PH2 dataset**

| | PSOVA | BA | DA | HS | FPA | MFO | ABC | CA | GA | PSO |
|---|---|---|---|---|---|---|---|---|---|---|
| **Average no. of selected features** | 50.23 | 68.60 | 69.60 | 65.33 | 64.83 | 57.70 | 66.64 | 61.37 | 62.20 | 66.63 |
| **Training cost (seconds)** | 3897.92 | 3808.29 | 3806.97 | 3927.84 | 3798.13 | 3900.98 | 3897.18 | 3922.39 | 3782.29 | 3876.96 |

| | PSOVA | BBPSO | ThBPSO | BBPSOV | AGPSO | ELPSO | DNLPSO | GM-PSO | MSPSO | GPSO |
|---|---|---|---|---|---|---|---|---|---|---|
| **Average no. of selected features** | 50.23 | 64.67 | 48.53 | 61.37 | 64.67 | 67.80 | 64.60 | 65.50 | 58.90 | 63.00 |
| **Training cost (seconds)** | 3897.92 | 3931.76 | 4039.71 | 3975.86 | 3877.59 | 3962.14 | 3855.17 | 3874.13 | 3997.29 | 3940.21 |

### 5.2.4 Evaluation Using the UCI Spam Base Mail Dataset

An additional larger scale UCI dataset was also used for evaluating the performance of feature optimisation for the proposed PSO algorithm. Specifically, the UCI spam base dataset (Mark Hopkins, 1998) was employed as part of the evaluation procedure to measure the discriminative abilities of the proposed model. The reason of using this dataset is because it has larger sample sizes than those of the other used medical datasets from UCI. Therefore, it is extracted and used. The spam base dataset consists of 4601 samples with 57 features, and two classes (positive and negative). All samples were used in the evaluation with an 80:20 aspect ratio of training and testing. In total, 30 runs were conducted for every method, with 57 dimensions, 20 populations and 200 iterations as the experimental settings. All methods utilised a maximum number of function evaluations, i.e. 20 population $\times$ 200 maximum number of iterations. The complete averages of GM scores over 30 runs for every other method are shown in Table 5.11 with the Wilcoxon rank sum test results also shown in Table 5.12. These results highlighted the significant performance improvements of the proposed PSO model against those of other alternative methods; the only exception was GM-PSO, which shows the same result distributions in comparison with those of the proposed model when integrated with the KNN classifier. Overall, the statistical findings demonstrate the superiority of the proposed PSO model when compared with alternative methods for deriving solutions relating to discriminative feature selection concerning this spam base data set.

**Table 5.11: Average performance of each algorithm over 30 runs for the spam base dataset**

| | KNN | | SVM | |
|---|---|---|---|---|
| | **10-fold** | **Hold-out** | **10-fold** | **Hold-out** |
| **BA** | 0.8645 | 0.8652 | 0.8810 | 0.8817 |
| **DA** | 0.8727 | 0.8734 | 0.8832 | 0.8838 |
| **HS** | 0.8809 | 0.8815 | 0.8836 | 0.8843 |
| **FPA** | 0.8724 | 0.8731 | 0.8883 | 0.8889 |
| **MFO** | 0.8809 | 0.8817 | 0.9009 | 0.9014 |
| **ABC** | 0.8916 | 0.8922 | 0.8986 | 0.8991 |
| **CA** | 0.8807 | 0.8816 | 0.9002 | 0.9008 |
| **GA** | 0.8757 | 0.8765 | 0.8911 | 0.8919 |
| **PSO** | 0.8762 | 0.8769 | 0.8887 | 0.8893 |
| **BBPSO** | 0.8799 | 0.8806 | 0.8930 | 0.8936 |
| **ThBPSO** | 0.8767 | 0.8774 | 0.8919 | 0.8925 |
| **BBPSOV** | 0.8726 | 0.8733 | 0.8939 | 0.8945 |
| **AGPSO** | 0.8817 | 0.8824 | 0.8976 | 0.8982 |
| **ELPSO** | 0.8714 | 0.8721 | 0.8679 | 0.8686 |
| **DNLPSO** | 0.8758 | 0.8764 | 0.8948 | 0.8954 |
| **GM-PSO** | 0.8711 | 0.8718 | 0.8850 | 0.8856 |
| **MS-PSO** | 0.8820 | 0.8827 | 0.8989 | 0.8995 |
| **GPSO** | 0.8669 | 0.8678 | 0.8792 | 0.8799 |

**Table 5.12: The *p-values* of the Wilcoxon rank sum test using the spam base dataset**

| | **BA** | **DA** | **HS** | **FPA** | **MFO** | **ABC** | **CA** | **GA** | **PSO** |
|---|---|---|---|---|---|---|---|---|---|
| **KNN Wilcoxon 10F** | 1.70E-08 | 7.22E-06 | 2.61E-02 | 3.57E-06 | 1.76E-02 | 3.92E-02 | 4.03E-03 | 2.77E-05 | 4.64E-05 |
| **KNN Wilcoxon HO** | 1.43E-08 | 6.49E-06 | 2.97E-02 | 2.15E-06 | 1.88E-02 | 3.78E-02 | 3.76E-03 | 2.68E-05 | 3.15E-05 |
| **SVM Wilcoxon 10F** | 3.02E-11 | 3.02E-11 | 2.87E-10 | 3.02E-11 | 8.99E-11 | 2.61E-10 | 6.70E-11 | 3.02E-11 | 3.02E-11 |
| **SVM Wilcoxon HO** | 3.00E-11 | 3.32E-11 | 2.14E-10 | 3.00E-11 | 6.61E-11 | 4.15E-10 | 6.65E-11 | 3.00E-11 | 3.00E-11 |

| | **BBPSO** | **ThBPSO** | **BBPSOV** | **AGPSO** | **ELPSO** | **DNLPSO** | **GM-PSO** | **MSPSO** | **GPSO** |
|---|---|---|---|---|---|---|---|---|---|
| **KNN Wilcoxon 10F** | 1.76E-02 | 4.51E-08 | 4.24E-02 | 1.49E-04 | 2.53E-04 | 7.73E-03 | **6.15E-02** | 1.29E-09 | 8.88E-06 |
| **KNN Wilcoxon HO** | 1.59E-02 | 2.86E-08 | 3.35E-02 | 1.36E-04 | 3.18E-04 | 6.78E-03 | **5.84E-02** | 1.07E-09 | 4.90E-06 |
| **SVM Wilcoxon 10F** | 6.70E-11 | 3.03E-09 | 7.39E-11 | 3.02E-11 | 1.86E-09 | 3.34E-11 | 2.37E-10 | 3.02E-11 | 3.02E-11 |
| **SVM Wilcoxon HO** | 6.66E-11 | 2.74E-09 | 7.34E-11 | 2.99E-11 | 1.69E-09 | 3.32E-11 | 2.13E-10 | 3.00E-11 | 3.00E-11 |

**Table 5.13: The training computational cost and the average number of features selected using the spam base dataset**

| | **PSOVA** | **BA** | **DA** | **HS** | **FPA** | **MFO** | **ABC** | **CA** | **GA** | **PSO** |
|---|---|---|---|---|---|---|---|---|---|---|
| **Average no. of selected features** | 33.73 | 36.43 | 3610 | 45.13 | 38.80 | 51.43 | 50.97 | 39.40 | 38.70 | 41.60 |
| **Training cost (seconds)** | 6916.25 | 6891.09 | 6811.01 | 6963.88 | 6858.62 | 6995.37 | 6947.75 | 6934.96 | 6793.35 | 6946.64 |

| | **PSOVA** | **BBPSO** | **ThBPSO** | **BBPSOV** | **AGPSO** | **ELPSO** | **DNLPSO** | **GM-PSO** | **MSPSO** | **GPSO** |
|---|---|---|---|---|---|---|---|---|---|---|
| **Average no. of selected features** | 33.73 | 43.90 | 32.98 | 47.33 | 45.20 | 47.57 | 43.73 | 48.20 | 31.93 | 43.70 |
| **Training cost (seconds)** | 6916.25 | 7005.57 | 7136.04 | 6899.99 | 6891.06 | 7037.60 | 6947.66 | 6957.73 | 6983.78 | 7026.52 |

### 5.2.5 Evaluation Using Benchmark Functions

The proposed PSOVA model is designed for maximisation with the features selection problem and achieves a superior performance, as shown in sections 0-5.2.3. Additionally, the proposed algorithm is also tested with the minimisation problem; therefore, the test

function has been used for further examination. The same benchmark functions in Table 4.13 have also been used here, i.e. Dixon-Price, Sphere, Rotated Hyper-Ellipsoid, Sum Squares, Sum of Different Powers, Ackley, Griewank and Powell, for evaluation of the PSOVA model.

## I. Results for Benchmark Functions with Low Dimension ($D = 10$)

The experiment begins with low-dimension minimization problems. The proposed PSOVA model has the most optimal results (close to 0) in comparison with those of other optimisation algorithms for nearly all benchmark functions. The performance of the PSO algorithm also exceeds those of other optimisation algorithms statistically in nearly all the cases. The efficiency of search space exploration is enhanced by the proposed PSO model underpinned by sequential sub-swarms, full and partial avoidance and attraction actions, thus substantially alleviating the problem of premature convergence (see Table 5.14). The performance of the proposed PSO model is only slightly lower compared to that of AGPSO for the Powell function. As shown Figure 5-1, the proposed model shows fast convergence rate in most of the cases in comparison with those of other methods. Table 5.15 lists the statistical test results, i.e. the *p-values* from the Wilcoxon rank sum test associated with this dimension. These values indicate that the results of PSOVA are significantly better than those of nearly all other optimisation algorithms. However, it must be highlighted that, despite having a higher performance compared to those of CA, ELPSO and DNLPSO for Ackley, PSOVA does not show sufficiently different mean GM performances in comparison with those of these three methods.

**Table 5.14 : Evaluation results for benchmark functions with $D = 10$**

| | | PSOVA | BA | DA | HS | FPA | MFO | ABC | CA | GA |
|---|---|---|---|---|---|---|---|---|---|---|
| **Dixon-Price** | **Mean** | **4.44E-01** | 4.55E+01 | 9.61E+00 | 6.14E-01 | 4.85E+00 | 3.56E+01 | 6.67E-01 | 7.17E-01 | 1.78E+03 |
| | **Std** | 3.20E-01 | 7.56E+01 | 2.87E+01 | 3.35E-01 | 2.01E+00 | 8.36E+01 | 1.74E-03 | 1.91E-01 | 2.15E+03 |
| | **Min** | 9.62E-22 | 6.67E-01 | 6.67E-01 | 1.80E-02 | 1.96E+00 | 3.64E-09 | 6.67E-01 | 3.38E-03 | 8.27E+01 |
| | **Max** | **6.67E-01** | 3.42E+02 | 1.24E+02 | 1.01E+00 | 9.30E+00 | 3.21E+02 | 6.76E-01 | 1.18E+00 | 9.76E+03 |
| **Sphere** | **Mean** | **6.45E-37** | 1.05E-06 | 3.09E-03 | 1.05E-04 | 1.26E-01 | 2.47E-16 | 2.90E-11 | 2.50E-28 | 1.71E+00 |
| | **Std** | **1.72E-36** | 2.19E-07 | 5.27E-03 | 5.05E-05 | 5.60E-02 | 4.10E-16 | 3.86E-11 | 1.21E-27 | 1.13E+00 |
| | **Min** | 6.13E-40 | 6.37E-07 | **0.00E+00** | 2.87E-05 | 4.35E-02 | 6.67E-19 | 1.56E-12 | 1.33E-52 | 3.96E-01 |
| | **Max** | **8.62E-36** | 1.61E-06 | 1.85E-02 | 2.69E-04 | 3.24E-01 | 1.54E-15 | 1.53E-10 | 6.62E-27 | 5.52E+00 |
| **Rotated Hyper-Ellipsoid** | **Mean** | **6.42E-34** | 5.21E+03 | 2.82E+01 | 5.83E-04 | 8.69E+01 | 1.43E+02 | 1.30E-08 | 7.73E-29 | 3.54E+03 |
| | **Std** | **2.45E-33** | 2.27E+03 | 8.02E+01 | 2.65E-04 | 3.23E+01 | 7.84E+02 | 1.28E-08 | 4.23E-28 | 2.88E+03 |
| | **Min** | 1.08E-37 | 1.76E+03 | 1.08E-30 | 1.34E-04 | 9.08E+00 | 2.49E-16 | 1.01E-09 | 9.14E-49 | 9.99E+02 |
| | **Max** | **1.35E-32** | 1.04E+04 | 4.05E+02 | 1.26E-03 | 1.34E+02 | 4.29E+03 | 5.43E-08 | 2.32E-27 | 1.61E+04 |
| **Sum Squares** | **Mean** | **1.93E-36** | 8.16E+00 | 3.31E-01 | 4.28E-04 | 2.33E+00 | 2.82E-15 | 5.51E-10 | 8.26E-35 | 4.87E+01 |
| | **Std** | 8.37E-35 | 1.15E+01 | 5.33E-01 | 1.73E-04 | 9.11E-01 | 4.56E-15 | 1.07E-09 | 4.45E-34 | 2.29E+01 |
| | **Min** | 8.00E-40 | 6.18E-06 | **0.00E+00** | 1.73E-04 | 9.95E-01 | 9.50E-18 | 1.79E-11 | 2.71E-51 | 1.82E+01 |
| | **Max** | **4.62E-34** | 4.84E+01 | 2.67E+00 | 7.78E-04 | 5.47E+00 | 1.58E-14 | 4.42E-09 | 2.44E-33 | 9.76E+01 |
| **Sum of Different** | **Mean** | **7.54E-62** | 1.33E-08 | 4.33E-06 | 3.56E-08 | 4.77E-06 | 8.10E-35 | 1.66E-17 | 9.48E-11 | 1.03E-02 |
| | **Std** | **2.27E-61** | 6.23E-09 | 9.72E-06 | 4.84E-08 | 4.03E-06 | 3.42E-34 | 3.16E-17 | 2.96E-10 | 1.46E-02 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Powers** | **Min** | **8.62E-68** | 1.33E-09 | 2.73E-16 | 7.82E-12 | 5.54E-07 | 3.32E-42 | 4.02E-20 | 1.60E-36 | 3.39E-05 |
| | **Max** | **8.47E-61** | 2.54E-08 | 5.17E-05 | 2.05E-07 | 1.89E-05 | 1.86E-33 | 1.43E-16 | 1.27E-09 | 6.26E-02 |
| **Ackley** | **Mean** | **7.40E-15** | 1.38E+01 | 2.02E+00 | 2.08E-02 | 6.46E+00 | 4.87E-08 | 6.21E-04 | 3.48E-01 | 1.13E+01 |
| | **Std** | **2.10E-15** | 1.97E+00 | 1.18E+00 | 7.78E-03 | 8.34E-01 | 3.98E-08 | 3.68E-04 | 7.17E-01 | 1.90E+00 |
| | **Min** | **4.44E-15** | 9.00E+00 | **4.44E-15** | 8.30E-03 | 4.35E+00 | 4.63E-09 | 1.67E-04 | **4.44E-15** | 6.99E+00 |
| | **Max** | **1.51E-14** | 1.67E+01 | 5.38E+00 | 4.19E-02 | 7.77E+00 | 1.74E-07 | 1.68E-03 | 3.03E+00 | 1.44E+01 |
| **Griewank** | **Mean** | **7.58E-02** | 2.89E+01 | 4.19E-01 | 1.00E-01 | 1.38E+00 | 1.72E-01 | 3.11E-01 | 1.49E-01 | 1.74E+01 |
| | **Std** | **2.61E-02** | 1.57E+01 | 3.62E-01 | 5.37E-02 | 1.32E-01 | 9.40E-02 | 6.43E-02 | 7.58E-02 | 8.02E+00 |
| | **Min** | 2.96E-02 | 1.23E+01 | **0.00E+00** | 7.43E-03 | 1.07E+00 | 3.94E-02 | 1.64E-01 | 5.17E-02 | 3.80E+00 |
| | **Max** | **1.38E-01** | 7.14E+01 | 1.57E+00 | 2.34E-01 | 1.60E+00 | 4.50E-01 | 4.66E-01 | 3.45E-01 | 3.32E+01 |
| **Powell** | **Mean** | 1.24E-04 | 1.14E-01 | 1.01E+01 | 8.04E-03 | 2.53E-02 | 2.12E+00 | 2.14E-04 | 7.94E-03 | 3.92E+01 |
| | **Std** | 9.09E-05 | 5.26E-01 | 2.41E+01 | 4.37E-03 | 1.24E-02 | 1.16E+01 | 1.28E-04 | 2.09E-02 | 2.89E+01 |
| | **Min** | 1.47E-05 | 1.27E-04 | 3.64E-03 | 2.85E-04 | 6.48E-03 | 1.95E-05 | 5.73E-05 | 2.17E-05 | 8.18E+00 |
| | **Max** | 3.71E-04 | 2.88E+00 | 7.73E+01 | 1.99E-02 | 5.85E-02 | 6.36E+01 | 5.20E-04 | 1.14E-01 | 1.19E+02 |

| | | **PSOVA** | **PSO** | **BBPSO** | **BBPSOV** | **AGPSO** | **ELPSO** | **DNLPSO** | **MSPSO** | **GPSO** |
|---|---|---|---|---|---|---|---|---|---|---|
| **Dixon-Price** | **Mean** | **4.44E-01** | 5.31E+03 | 4.71E+01 | 6.66E-01 | 5.56E-01 | 1.00E+01 | 6.59E-01 | 1.99E+01 | 4.76E+03 |
| | **Std** | 3.20E-01 | 3.35E+03 | 9.85E+01 | 3.46E-01 | 2.53E-01 | **0.00E+00** | 1.35E-01 | 2.07E+01 | 2.53E+03 |
| | **Min** | 9.62E-22 | 8.28E+02 | 1.29E-20 | 3.75E-06 | 2.86E-23 | 1.00E+01 | **1.37E-30** | 1.07E+00 | 1.78E+03 |
| | **Max** | **6.67E-01** | 1.59E+04 | 3.21E+02 | 1.47E+00 | **6.67E-01** | 1.00E+01 | 8.89E-01 | 7.79E+01 | 1.26E+04 |
| **Sphere** | **Mean** | **6.45E-37** | 1.11E+01 | 8.70E-36 | 4.24E-07 | 3.89E-32 | 4.61E+00 | 3.07E-02 | 2.65E-01 | 9.21E+00 |
| | **Std** | **1.72E-36** | 3.25E+00 | 4.40E-35 | 1.82E-06 | 2.11E-31 | 1.55E+00 | 1.39E-01 | 2.66E-01 | 2.52E+00 |
| | **Min** | 6.13E-40 | 3.54E+00 | 5.30E-41 | 3.06E-24 | 1.02E-37 | 2.89E-12 | 2.26E-58 | 3.27E-03 | 4.75E+00 |
| | **Max** | **8.62E-36** | 1.83E+01 | 2.41E-34 | 9.83E-06 | 1.15E-30 | 5.12E+00 | 7.49E-01 | 9.62E-01 | 1.39E+01 |
| **Rotated Hyper-Ellipsoid** | **Mean** | **6.42E-34** | 9.06E+03 | 1.43E+02 | 1.54E+00 | 1.62E-24 | 6.55E+01 | 3.33E-01 | 2.82E+02 | 7.52E+03 |
| | **Std** | **2.45E-33** | 3.20E+03 | 7.84E+02 | 8.34E+00 | 8.87E-24 | 4.34E-14 | 1.80E+00 | 2.09E+02 | 2.76E+03 |
| | **Min** | 1.08E-37 | 2.12E+03 | 3.02E-38 | 1.84E-17 | 7.07E-35 | 6.55E+01 | **3.46E-53** | 2.35E+01 | 3.53E+03 |
| | **Max** | **1.35E-32** | 1.71E+04 | 4.29E+03 | 4.57E+01 | 4.86E-23 | 6.55E+01 | 9.88E+00 | 7.62E+02 | 1.49E+04 |
| **Sum Squares** | **Mean** | **1.93E-36** | 1.79E+02 | 1.36E-35 | 6.11E-06 | 4.09E-30 | 1.00E+01 | 5.42E-03 | 6.38E+00 | 1.70E+02 |
| | **Std** | 8.37E-35 | 6.45E+01 | 3.14E-35 | 3.07E-05 | 2.17E-29 | **0.00E+00** | 2.97E-02 | 4.92E+00 | 5.02E+01 |
| | **Min** | 8.00E-40 | 6.76E+01 | 2.78E-40 | 2.31E-20 | 5.45E-36 | 1.00E+01 | 7.83E-56 | 3.62E-01 | 5.54E+01 |
| | **Max** | **4.62E-34** | 3.26E+02 | 1.38E-33 | 1.68E-04 | 1.19E-28 | 1.00E+01 | 1.62E-01 | 2.01E+01 | 2.60E+02 |
| **Sum of Different Powers** | **Mean** | **7.54E-62** | 1.88E-02 | 1.27E-52 | 5.34E-20 | 3.32E-43 | 7.32E-56 | 8.88E-11 | 2.29E-05 | 1.42E-02 |
| | **Std** | **2.27E-61** | 1.02E-02 | 6.79E-62 | 2.13E-19 | 1.43E-42 | 3.50E-55 | 4.13E-10 | 4.05E-05 | 7.87E-03 |
| | **Min** | **8.62E-68** | 2.93E-03 | 9.76E-74 | 5.40E-42 | 5.17E-51 | 2.88E-66 | 4.41E-67 | 6.93E-07 | 3.39E-03 |
| | **Max** | **8.47E-61** | 4.08E-02 | 3.72E-61 | 1.11E-18 | 7.83E-42 | 1.92E-54 | 2.25E-09 | 2.01E-04 | 3.39E-02 |
| **Ackley** | **Mean** | **7.40E-15** | 1.64E+01 | 1.45E+00 | 1.31E+00 | 5.28E-10 | 8.23E-15 | 6.82E-01 | 5.43E+00 | 1.53E+01 |
| | **Std** | **2.10E-15** | 1.22E+00 | 4.84E+00 | 1.49E+00 | 2.89E-09 | 2.63E-15 | 3.65E+00 | 1.72E+00 | 1.46E+00 |
| | **Min** | **4.44E-15** | 1.34E+01 | **4.44E-15** | 1.90E-11 | **4.44E-15** | **4.44E-15** | **4.44E-15** | 1.66E+00 | 1.10E+01 |
| | **Max** | **1.51E-14** | 1.81E+01 | 1.94E+01 | 4.77E+00 | 1.58E-08 | **1.51E-14** | 2.00E+01 | 8.39E+00 | 1.78E+01 |
| **Griewank** | **Mean** | **7.58E-02** | 4.10E+01 | 1.39E-01 | 2.76E-01 | 8.79E-02 | 8.56E-02 | 1.29E-01 | 1.72E+00 | 3.23E+01 |
| | **Std** | **2.61E-02** | 9.48E+00 | 8.35E-02 | 7.81E-01 | 5.36E-02 | 4.11E-02 | 9.28E-02 | 1.15E+00 | 8.69E+00 |
| | **Min** | 2.96E-02 | 2.26E+01 | 3.20E-02 | 1.48E-02 | 1.48E-02 | 2.71E-02 | 2.13E-02 | 4.83E-01 | 1.59E+01 |
| | **Max** | **1.38E-01** | 5.36E+01 | 3.37E-01 | 4.39E+00 | 2.39E-01 | 1.87E-01 | 4.53E-01 | 5.19E+00 | 5.47E+01 |
| **Powell** | **Mean** | 1.24E-04 | 1.08E+02 | 8.48E+00 | 6.36E-03 | **4.41E-06** | 5.00E+00 | 1.36E-04 | 2.91E+00 | 9.96E+01 |
| | **Std** | 9.09E-05 | 4.06E+01 | 2.62E+01 | 1.53E-02 | 2.56E-06 | **0.00E+00** | 4.87E-04 | 3.16E+00 | 3.90E+01 |
| | **Min** | 1.47E-05 | 4.30E+01 | 8.85E-07 | 3.70E-05 | 3.19E-07 | 5.00E+00 | **1.18E-07** | 7.02E-04 | 2.38E+01 |
| | **Max** | 3.71E-04 | 2.26E+02 | 1.04E+02 | 7.79E-02 | **1.09E-05** | 5.00E+00 | 2.65E-03 | 1.19E+01 | 1.70E+02 |

**Table 5.15: The Wilcoxon rank sum test results for all benchmark functions with $D$ = 10**

| | **BA** | **DA** | **HS** | **FPA** | **MFO** | **ABC** | **CA** | **GA** |
|---|---|---|---|---|---|---|---|---|
| **Dixon** | 8.08E-11 | 2.99E-11 | 5.25E-05 | 2.99E-11 | 8.76E-03 | 1.94E-10 | 1.19E-03 | 2.99E-11 |
| **Sphere** | 3.02E-11 | 8.48E-09 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 8.48E-09 | 3.02E-11 |
| **Rotated Hyper-Ellipsoid** | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 1.17E-09 | 3.02E-11 |
| **Sum Squares** | 3.02E-11 | 8.48E-09 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 1.31E-08 | 3.02E-11 |
| **Sum of Different Powers** | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 |
| **Ackley** | 1.69E-11 | 1.89E-10 | 1.69E-11 | 1.69E-11 | 1.69E-11 | 1.69E-11 | **2.39E-01** | 1.69E-11 |
| **Griewank** | 3.02E-11 | 1.36E-07 | 3.18E-04 | 3.02E-11 | 5.00E-09 | 3.02E-11 | 7.12E-09 | 3.02E-11 |
| **Powell** | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 |

| | **PSO** | **BBPSO** | **BBPSOV** | **AGPSO** | **ELPSO** | **DNLPSO** | **MSPSO** | **GPSO** |
|---|---|---|---|---|---|---|---|---|
| **Dixon** | 2.99E-11 | 5.28E-03 | 3.83E-02 | 1.59E-02 | 1.20E-12 | 3.06E-04 | 2.99E-11 | 2.99E-11 |
| **Sphere** | 3.02E-11 | 2.46E-03 | 3.02E-11 | 5.09E-06 | 3.16E-12 | 1.00E+03 | 3.02E-11 | 3.02E-11 |
| **Rotated Hyper-Ellipsoid** | 3.02E-11 | 3.95E-04 | 3.02E-11 | 4.31E-08 | 1.21E-12 | 3.87E-03 | 3.02E-11 | 3.02E-11 |
| **Sum Squares** | 3.02E-11 | 9.71E-03 | 3.02E-11 | 3.20E-09 | 1.21E-12 | 3.18E-03 | 3.02E-11 | 3.02E-11 |
| **Sum of Different Powers** | 3.02E-11 | 2.12E-04 | 3.02E-11 | 3.02E-11 | 4.08E-11 | 3.69E-11 | 3.02E-11 | 3.02E-11 |
| **Ackley** | 1.69E-11 | 1.56E-04 | 1.69E-11 | 5.78E-03 | **9.88E-02** | **4.15E-01** | 1.69E-11 | 1.69E-11 |
| **Griewank** | 3.02E-11 | 2.13E-05 | 1.34E-05 | 1.99E-02 | 2.62E-03 | 1.41E-04 | 3.02E-11 | 3.02E-11 |
| **Powell** | 3.02E-11 | 2.87E-10 | 3.02E-11 | 1.86E-09 | 1.21E-12 | 7.22E-06 | 3.02E-11 | 3.02E-11 |

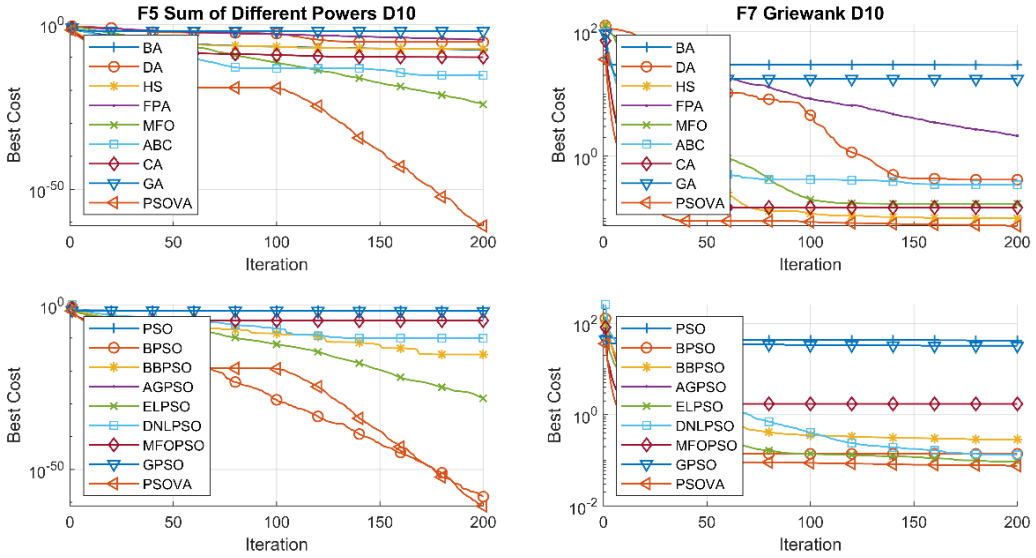



**Figure 5-1: Convergence curves for F2 and F6 in $D$ = 10**

## II. Results for Benchmark Functions with Medium Dimension ($D$ = 30)

Regarding the medium dimension, the same experimental setting was used for all algorithms, and the established results of the algorithms regarding the unimodal and multimodal functions are shown in Table 5.12. As shown in Figure 5-2, the proposed PSOVA model has the best convergence rate compared with those of other methods. Specifically, Figure 5-2 clearly indicates that the PSOVA model has great efficiency in search exploration, with fast convergence rates for F5 and F7, although its performance on the Griewank test function is lower than that of the CA and ELPSO. Furthermore, unlike other algorithms that produce suboptimal results, PSOVA is capable of detecting a near-global solution. Table 5.17 lists the *p-values* from the Wilcoxon rank sum test associated with this dimension. These values indicate that the results of PSOVA are optimistic and statistically better in comparison with those of nearly all other optimisation algorithms. As shown in this table, the proposed model outperforms all other methods for nearly all the benchmark functions, except for Griewank, where AGPSO and ELPSO show similar mean GM performances to those of the proposed model.

**Table 5.16 : Evaluation results for benchmark functions with $D$ = 30**

| | | PSOVA | BA | DA | HS | FPA | MFO | ABC | CA | GA |
|---|---|---|---|---|---|---|---|---|---|---|
| Dixon-Price | Mean | **7.51E-01** | 1.58E+04 | 4.18E+02 | 4.86E+00 | 5.50E+03 | 4.37E+04 | 1.58E+03 | 6.85E+01 | 7.42E+04 |
| | Std | 4.56E-01 | 1.71E+04 | 8.03E+02 | 1.78E+00 | 2.66E+03 | 8.99E+04 | 7.83E+02 | 9.79E+01 | 4.23E+04 |
| | Min | **6.67E-01** | 2.88E+02 | 1.21E+01 | 2.51E+00 | 2.17E+03 | 3.70E+00 | 6.40E+02 | 6.58E+00 | 1.95E+04 |
| | Max | **3.16E+00** | 7.10E+04 | 4.43E+03 | 9.08E+00 | 1.26E+04 | 3.46E+05 | 3.81E+03 | 4.73E+02 | 1.54E+05 |
| Sphere | Mean | **2.09E-21** | 2.43E+00 | 2.56E+00 | 6.00E-02 | 7.79E+00 | 4.37E+00 | 2.65E-02 | 7.03E-03 | 1.69E+01 |
| | Std | **4.04E-21** | 2.41E+00 | 2.19E+00 | 7.48E-03 | 2.33E+00 | 9.94E+00 | 8.87E-03 | 1.04E-02 | 8.55E+00 |
| | Min | **3.00E-24** | 1.61E-05 | 3.14E-01 | 5.00E-02 | 2.29E+00 | 5.48E-04 | 1.19E-02 | 9.21E-09 | 6.95E+00 |
| | Max | **1.92E-20** | 8.73E+00 | 7.81E+00 | 7.95E-02 | 1.39E+01 | 2.62E+01 | 4.97E-02 | 4.74E-02 | 4.81E+01 |
| Rotated Hyper-Ellipsoid | Mean | **3.60E-16** | 8.34E+04 | 5.08E+03 | 2.06E+00 | 1.66E+04 | 1.93E+04 | 4.14E+01 | 2.23E+01 | 7.43E+04 |
| | Std | **1.37E-15** | 2.82E+04 | 2.73E+03 | 5.83E-01 | 4.06E+03 | 1.89E+04 | 1.19E+01 | 4.68E+01 | 2.26E+04 |
| | Min | **2.16E-20** | 3.67E+04 | 6.71E+02 | 1.38E+00 | 8.93E+03 | 1.17E+00 | 2.35E+01 | 3.51E-03 | 4.18E+04 |
| | Max | **7.35E-15** | 1.41E+05 | 1.26E+04 | 4.48E+00 | 2.37E+04 | 6.44E+04 | 7.51E+01 | 2.20E+02 | 1.37E+05 |
| Sum Squares | Mean | **2.87E-18** | 6.71E+02 | 1.25E+02 | 9.39E-01 | 3.75E+02 | 5.10E+02 | 1.10E+00 | 1.37E+00 | 1.11E+03 |
| | Std | 9.23E-18 | 3.30E+02 | 1.13E+02 | 1.57E-01 | 9.80E+01 | 6.17E+02 | 2.93E-01 | 3.11E+00 | 3.50E+02 |
| | Min | **7.02E-23** | 7.65E+01 | 1.80E+00 | 5.66E-01 | 1.87E+02 | 4.40E-02 | 6.01E-01 | 2.38E-03 | 2.91E+02 |
| | Max | **4.48E-17** | 1.75E+03 | 4.88E+02 | 1.18E+00 | 6.06E+02 | 2.40E+03 | 1.82E+00 | 1.26E+01 | 1.83E+03 |
| Sum of Different Powers | Mean | **1.95E-33** | 6.43E-08 | 1.18E-05 | 2.41E-07 | 1.62E-04 | 3.07E-11 | 4.10E-03 | 5.20E-07 | 1.93E-02 |
| | Std | **9.42E-33** | 3.20E-08 | 2.42E-05 | 3.15E-07 | 1.60E-04 | 6.37E-11 | 3.17E-03 | 1.25E-06 | 2.21E-02 |
| | Min | 2.88E-44 | 1.76E-08 | 3.74E-10 | 1.10E-09 | 1.10E-05 | 1.02E-14 | 2.63E-04 | 7.67E-10 | 8.19E-04 |
| | Max | **5.17E-32** | 1.73E-07 | 1.32E-04 | 1.48E-06 | 6.97E-04 | 2.51E-10 | 1.52E-02 | 6.35E-06 | 8.77E-02 |
| Ackley | Mean | 2.03E+00 | 1.52E+01 | 7.76E+00 | 2.18E+00 | 1.31E+01 | 4.77E+00 | 3.85E+00 | 4.19E+00 | 1.45E+01 |
| | Std | 6.57E-01 | 1.16E+00 | 1.91E+00 | **4.68E-01** | 9.53E-01 | 6.75E+00 | 8.62E-01 | 1.23E+00 | 1.19E+00 |
| | Min | **1.36E-07** | 1.28E+01 | 2.06E+00 | 6.00E-01 | 1.07E+01 | 2.20E-01 | 2.64E+00 | 2.32E+00 | 1.25E+01 |
| | Max | 3.34E+00 | 1.71E+01 | 1.09E+01 | 2.79E+00 | 1.47E+01 | 1.91E+01 | 6.65E+00 | 6.68E+00 | 1.69E+01 |
| Griewank | Mean | 1.95E-02 | 1.50E+02 | 1.02E+01 | 1.09E+00 | 2.95E+01 | 2.49E+01 | 1.07E+00 | 4.91E-01 | 1.11E+02 |
| | Std | 2.31E-02 | 5.20E+01 | 7.45E+00 | 3.66E-02 | 8.26E+00 | 4.69E+01 | 3.72E-02 | 4.91E-01 | 3.36E+01 |
| | Min | **0.00E+00** | 6.82E+01 | 2.77E+00 | 1.04E+00 | 1.64E+01 | 4.40E-01 | 1.04E+00 | 2.47E-02 | 6.50E+01 |
| | Max | **8.30E-02** | 2.69E+02 | 3.51E+01 | 1.17E+00 | 5.56E+01 | 1.81E+02 | 1.19E+00 | 1.88E+00 | 2.52E+02 |
| Powell | Mean | **2.61E-03** | 4.53E+01 | 5.34E+01 | 2.45E+00 | 9.57E+01 | 4.28E+02 | 6.60E+02 | 1.96E+01 | 6.40E+02 |
| | Std | 1.74E-03 | 8.09E+01 | 5.63E+01 | 1.10E+00 | 3.84E+01 | 7.36E+02 | 2.35E+02 | 1.84E+01 | 3.86E+02 |
| | Min | 6.61E-04 | 1.87E+00 | 3.90E+00 | 9.35E-01 | 3.62E+01 | 2.51E+00 | 3.02E+02 | 1.41E-01 | 1.93E+02 |
| | Max | **1.04E-02** | 4.14E+02 | 2.60E+02 | 6.06E+00 | 1.77E+02 | 3.27E+03 | 1.23E+03 | 6.62E+01 | 1.81E+03 |

| | | PSOVA | PSO | BBPSO | BBPSOV | AGPSO | ELPSO | DNLPSO | MSPSO | GPSO |
|---|---|---|---|---|---|---|---|---|---|---|
| Dixon-Price | Mean | **7.51E-01** | 6.92E+05 | 5.80E+04 | 9.66E+02 | 1.23E+01 | 1.00E+01 | 2.71E+02 | 2.17E+04 | 4.88E+05 |
| | Std | 4.56E-01 | 3.12E+05 | 9.26E+04 | 1.47E+03 | 5.83E+01 | **0.00E+00** | 1.16E+03 | 1.02E+04 | 2.10E+05 |
| | Min | **6.67E-01** | 2.32E+05 | 6.81E-01 | 2.11E+01 | 6.67E-01 | 1.00E+01 | 6.67E-01 | 9.45E+03 | 1.95E+05 |
| | Max | **3.16E+00** | 1.33E+06 | 2.74E+05 | 5.43E+03 | 3.21E+02 | 1.00E+01 | 6.14E+03 | 4.77E+04 | 1.19E+06 |
| Sphere | Mean | **2.09E-21** | 8.83E+01 | 9.61E+00 | 3.40E+00 | 5.97E-07 | 5.12E+00 | 4.44E-01 | 1.69E+01 | 8.50E+01 |
| | Std | **4.04E-21** | 2.17E+01 | 1.46E+01 | 3.60E+00 | 9.88E-07 | 1.81E-15 | 1.32E+00 | 3.12E+00 | 1.54E+01 |
| | Min | **3.00E-24** | 4.56E+01 | 1.06E-07 | 5.70E-03 | 4.23E-08 | 5.12E+00 | 1.84E-13 | 9.64E+00 | 5.49E+01 |
| | Max | **1.92E-20** | 1.43E+02 | 5.24E+01 | 1.45E+01 | 5.21E-06 | 5.12E+00 | 5.32E+00 | 2.11E+01 | 1.25E+02 |
| Rotated Hyper-Ellipsoid | Mean | **3.60E-16** | 2.01E+05 | 2.58E+04 | 8.19E+03 | 2.86E+02 | 6.55E+01 | 4.70E+02 | 3.62E+04 | 1.82E+05 |
| | Std | **1.37E-15** | 5.78E+04 | 2.22E+04 | 8.80E+03 | 1.09E+03 | 4.34E-14 | 1.46E+03 | 9.75E+03 | 3.25E+04 |
| | Min | **2.16E-20** | 1.04E+05 | 1.38E-03 | 8.11E+00 | 5.95E-05 | 6.55E+01 | 3.67E-12 | 1.76E+04 | 1.28E+05 |
| | Max | **7.35E-15** | 3.50E+05 | 7.73E+04 | 2.83E+04 | 4.29E+03 | 6.55E+01 | 7.35E+03 | 5.61E+04 | 2.65E+05 |
| Sum Squares | Mean | **2.87E-18** | 5.05E+03 | 5.57E+02 | 1.63E+02 | 3.33E+00 | 1.00E+01 | 2.18E-01 | 7.85E+02 | 4.25E+03 |
| | Std | 9.23E-18 | 1.28E+03 | 6.65E+02 | 1.82E+02 | 1.83E+01 | **0.00E+00** | 1.12E+00 | 2.11E+02 | 9.00E+02 |
| | Min | **7.02E-23** | 2.85E+03 | 3.82E-06 | 2.59E-01 | 1.10E-06 | 1.00E+01 | 8.67E-16 | 3.74E+02 | 2.55E+03 |
| | Max | **4.48E-17** | 8.15E+03 | 2.70E+03 | 8.76E+02 | 1.00E+02 | 1.00E+01 | 6.12E+00 | 1.18E+03 | 5.96E+03 |
| Sum of Different Powers | Mean | **1.95E-33** | 2.33E-01 | 3.62E-18 | 2.58E-09 | 3.85E-18 | 8.17E-14 | 1.16E-08 | 9.65E-05 | 1.57E-01 |
| | Std | **9.42E-33** | 2.03E-01 | 1.28E-17 | 5.97E-09 | 1.24E-17 | 4.29E-13 | 4.50E-08 | 1.64E-04 | 1.81E-01 |
| | Min | 2.88E-44 | 2.03E-02 | 4.22E-26 | 5.29E-16 | 4.94E-22 | 1.87E-22 | **4.22E-49** | 2.11E-06 | 1.35E-02 |
| | Max | **5.17E-32** | 6.37E-01 | 5.08E-17 | 2.48E-08 | 6.39E-17 | 2.35E-12 | 2.37E-07 | 8.18E-04 | 9.55E-01 |
| Ackley | Mean | 2.03E+00 | 1.94E+01 | 1.50E+01 | 8.57E+00 | 2.64E+00 | **2.65E-01** | 3.92E+00 | 1.36E+01 | 1.93E+01 |
| | Std | 6.57E-01 | 6.05E-01 | 5.60E+00 | 2.87E+00 | 7.42E-01 | 5.15E-01 | 3.05E+00 | 1.04E+00 | 6.45E-01 |
| | Min | **1.36E-07** | 1.78E+01 | 6.75E-01 | 3.15E+00 | 1.67E+00 | 3.94E-05 | 4.86E-04 | 1.16E+01 | 1.72E+01 |
| | Max | 3.34E+00 | 2.02E+01 | 1.94E+01 | 1.40E+01 | 4.89E+00 | **1.52E+00** | 1.40E+01 | 1.54E+01 | 2.02E+01 |
| Griewank | Mean | 1.95E-02 | 2.74E+02 | 1.84E+01 | 1.27E+01 | 2.91E-02 | **1.82E-02** | 3.85E+00 | 5.70E+01 | 2.84E+02 |
| | Std | 2.31E-02 | 7.92E+01 | 3.69E+01 | 1.03E+01 | 3.01E-02 | **2.03E-02** | 1.02E+01 | 1.07E+01 | 5.05E+01 |
| | Min | **0.00E+00** | 1.20E+02 | 8.53E-04 | 1.06E+00 | 2.17E-05 | 1.68E-06 | 1.69E-08 | 3.55E+01 | 1.79E+02 |
| | Max | **8.30E-02** | 4.26E+02 | 9.30E+01 | 4.31E+01 | 9.79E-02 | 8.37E-02 | 4.93E+01 | 7.96E+01 | 3.87E+02 |
| Powell | Mean | **2.61E-03** | 5.12E+03 | 6.44E+02 | 3.57E+01 | 1.48E+01 | 5.00E+00 | 1.94E+00 | 4.09E+02 | 4.05E+03 |
| | Std | 1.74E-03 | 1.26E+03 | 9.23E+02 | 3.44E+01 | 3.21E+01 | **0.00E+00** | 6.52E+00 | 1.77E+02 | 1.27E+03 |
| | Min | 6.61E-04 | 2.06E+03 | 1.35E-02 | 1.29E+00 | 8.15E-03 | 5.00E+00 | **1.93E-04** | 1.58E+02 | 1.49E+03 |
| | Max | **1.04E-02** | 7.91E+03 | 2.90E+03 | 1.62E+02 | 1.03E+02 | 5.00E+00 | 2.86E+01 | 7.52E+02 | 6.10E+03 |

**Table 5.17: The Wilcoxon rank sum test results for all benchmark functions with $D$ = 30**

| | BA | DA | HS | FPA | MFO | ABC | CA | GA |
|---|---|---|---|---|---|---|---|---|
| **Dixon** | 3.02E-11 | 3.02E-11 | 8.99E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 |
| **Sphere** | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 |
| **Rotated Hyper-Ellipsoid** | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 |
| **Sum Squares** | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 |
| **Sum of Different Powers** | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 |
| **Ackley** | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 |
| **Griewank** | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 7.38E-10 | 3.02E-11 |
| **Powell** | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 |

| | PSO | BBPSO | BBPSOV | AGPSO | ELPSO | DNLPSO | MSPSO | GPSO |
|---|---|---|---|---|---|---|---|---|
| **Dixon** | 3.02E-11 | 8.89E-10 | 3.02E-11 | 2.75E-03 | 1.21E-12 | 1.76E-04 | 3.02E-11 | 3.02E-11 |
| **Sphere** | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 1.21E-12 | 2.03E-07 | 3.02E-11 | 3.02E-11 |
| **Rotated Hyper-Ellipsoid** | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 1.21E-12 | 3.52E-07 | 3.02E-11 | 3.02E-11 |
| **Sum Squares** | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 1.21E-12 | 5.57E-03 | 3.02E-11 | 3.02E-11 |
| **Sum of Different Powers** | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 |
| **Ackley** | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 1.78E-10 | 3.69E-11 | 3.02E-11 | 3.02E-11 |
| **Griewank** | 3.02E-11 | 1.25E-07 | 3.02E-11 | **7.24E-02** | **5.11E-01** | 9.52E-04 | 3.02E-11 | 3.02E-11 |
| **Powell** | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 1.21E-12 | 5.94E-05 | 3.02E-11 | 3.02E-11 |

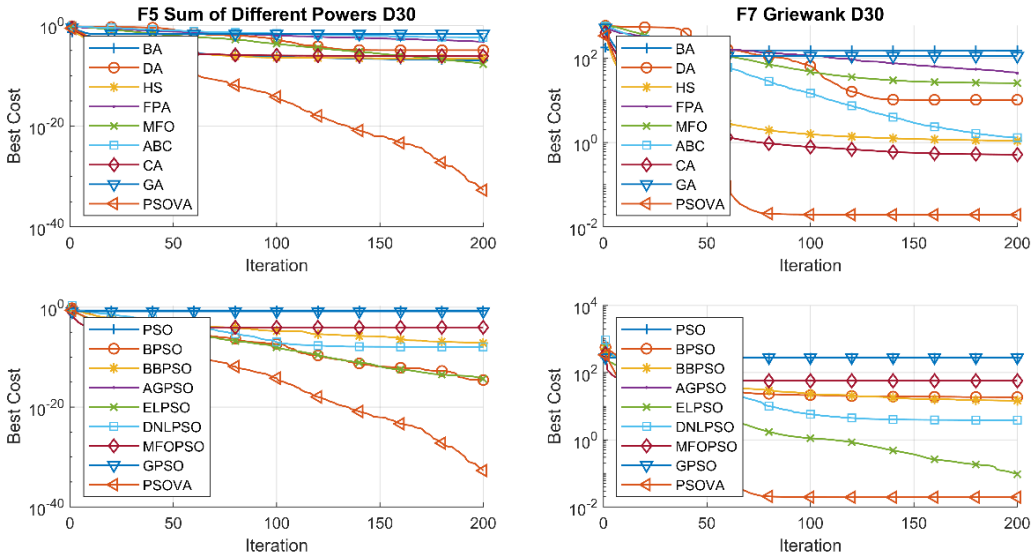



**Figure 5-2: Convergence curves for F5 and F7 in $D$ = 30**

## III. Results for Benchmark Functions with High Dimension ($D$ = 50)

The same experimental setting is also applied to the evaluation using dimension 50. Table 5.18 shows that PSOVA produces the most efficient results compared to the results of all other algorithms related to this scenario, except that ELPSO gains the best performance for the Ackley function. Moreover, the increase in dimensionality does not have a significant effect on the PSOVA outcomes for either unimodal or multimodal functions, which is a proof of the scalability of the proposed PSO optimisation algorithm. Figure 5.3 illustrates the fast convergence of the proposed model against those of other methods. The *p-values* obtained from the Wilcoxon rank sum test are provided in Table 5.19, which

clearly indicates the greater efficiency of the PSOVA algorithm compared to other algorithms, for solving high-dimensional optimisation problems. The exception is for Ackley, where AGPSO, DNLPSO and the proposed PSOVA model show similar result distributions.

**Table 5.18 : Evaluation results for benchmark functions with $D$ = 50**

| | | PSOVA | BA | DA | HS | FPA | MFO | ABC | CA | GA |
|---|---|---|---|---|---|---|---|---|---|---|
| Dixon-Price | Mean | 4.77E+00 | 6.71E+04 | 6.48E+03 | 7.91E+01 | 3.22E+04 | 2.84E+05 | 3.19E+06 | 4.96E+03 | 4.12E+05 |
| | Std | 3.16E+00 | 5.39E+04 | 6.75E+03 | 1.87E+01 | 1.27E+04 | 3.90E+05 | 8.10E+05 | 4.51E+03 | 2.49E+05 |
| | Min | 6.70E-01 | 2.89E+03 | 9.39E+01 | 4.54E+01 | 1.24E+04 | 8.59E+02 | 1.67E+06 | 5.91E+02 | 1.02E+05 |
| | Max | 9.82E+00 | 2.29E+05 | 2.42E+04 | 1.19E+02 | 6.55E+04 | 1.57E+06 | 4.97E+06 | 1.86E+04 | 1.01E+06 |
| Sphere | Mean | 5.21E-08 | 1.43E+01 | 9.91E+00 | 3.91E-01 | 1.83E+01 | 1.88E+01 | 1.54E+02 | 1.01E+00 | 4.14E+01 |
| | Std | 1.03E-07 | 9.78E+00 | 4.63E+00 | 4.28E-02 | 3.60E+00 | 1.97E+01 | 2.35E+01 | 1.25E+00 | 1.29E+01 |
| | Min | 9.30E-10 | 3.54E-01 | 3.73E+00 | 3.09E-01 | 1.19E+01 | 4.86E-01 | 1.06E+02 | 7.09E-02 | 2.31E+01 |
| | Max | 3.22E-07 | 3.33E+01 | 2.29E+01 | 4.61E-01 | 2.84E+01 | 7.92E+01 | 1.94E+02 | 6.42E+00 | 8.40E+01 |
| Rotated Hyper-Ellipsoid | Mean | 1.88E-04 | 2.75E+05 | 2.16E+04 | 1.68E+03 | 6.83E+04 | 1.11E+05 | 5.05E+05 | 4.85E+03 | 2.40E+05 |
| | Std | 3.94E-04 | 8.74E+04 | 1.39E+04 | 4.51E+02 | 1.49E+04 | 9.56E+04 | 7.24E+04 | 4.89E+03 | 4.88E+04 |
| | Min | 3.84E-06 | 1.08E+05 | 3.43E+03 | 7.65E+02 | 4.07E+04 | 9.17E+03 | 3.88E+05 | 5.01E+02 | 1.53E+05 |
| | Max | 1.74E-03 | 5.05E+05 | 5.79E+04 | 2.73E+03 | 1.01E+05 | 3.83E+05 | 6.54E+05 | 2.33E+04 | 3.87E+05 |
| Sum Squares | Mean | 4.78E-06 | 2.86E+03 | 7.16E+02 | 1.52E+01 | 1.66E+03 | 2.40E+03 | 1.18E+04 | 1.46E+02 | 4.09E+03 |
| | Std | 8.15E-06 | 9.49E+02 | 6.90E+02 | 2.16E+00 | 3.90E+02 | 2.35E+03 | 1.62E+03 | 1.91E+02 | 1.25E+03 |
| | Min | 1.41E-07 | 1.08E+03 | 2.35E+01 | 1.21E+01 | 9.81E+02 | 1.00E+02 | 8.60E+03 | 5.08E+00 | 2.44E+03 |
| | Max | 3.11E-05 | 4.99E+03 | 3.27E+03 | 2.27E+01 | 2.39E+03 | 9.02E+03 | 1.43E+04 | 1.01E+03 | 7.34E+03 |
| Sum of Different Powers | Mean | 3.28E-18 | 5.69E-08 | 6.73E-06 | 4.78E-07 | 2.94E-04 | 2.79E-06 | 7.59E-01 | 6.15E-05 | 2.39E-02 |
| | Std | 1.01E-17 | 2.95E-08 | 1.08E-05 | 6.68E-07 | 2.71E-04 | 4.67E-06 | 3.00E-01 | 1.38E-04 | 2.15E-02 |
| | Min | 1.20E-22 | 1.13E-08 | 2.52E-17 | 1.08E-08 | 9.78E-06 | 2.74E-08 | 1.87E-01 | 1.86E-07 | 2.92E-03 |
| | Max | 5.33E-17 | 1.17E-07 | 4.87E-05 | 3.45E-06 | 9.12E-04 | 1.70E-05 | 1.50E+00 | 6.89E-04 | 8.04E-02 |
| Ackley | Mean | 3.27E+00 | 1.57E+01 | 1.01E+01 | 3.09E+01 | 1.38E+01 | 1.85E+01 | 2.03E+01 | 1.10E+01 | 1.61E+01 |
| | Std | 8.23E-01 | 9.40E-01 | 2.03E+00 | 3.97E-01 | 9.09E-01 | 1.60E+00 | 2.54E-01 | 2.36E+00 | 8.56E-01 |
| | Min | 2.08E+00 | 1.37E+01 | 5.58E+00 | 3.02E+01 | 1.18E+01 | 1.31E+01 | 1.98E+01 | 5.46E+00 | 1.44E+01 |
| | Max | 5.68E+00 | 1.74E+01 | 1.38E+01 | 4.08E+01 | 1.53E+01 | 1.97E+01 | 2.08E+01 | 1.74E+01 | 1.81E+01 |
| Griewank | Mean | 1.78E-02 | 2.31E+02 | 2.95E+01 | 3.06E+00 | 6.38E+01 | 6.58E+01 | 5.59E+02 | 5.77E+00 | 2.56E+02 |
| | Std | 2.43E-02 | 6.55E+01 | 1.73E+01 | 4.47E-01 | 1.30E+01 | 6.33E+01 | 7.47E+01 | 5.48E+00 | 7.11E+01 |
| | Min | 1.14E-06 | 1.06E+02 | 6.04E+00 | 2.08E+00 | 3.96E+01 | 3.51E+00 | 3.73E+02 | 1.15E+00 | 1.45E+02 |
| | Max | 8.50E-02 | 4.14E+02 | 7.76E+01 | 3.78E+00 | 9.35E+01 | 1.87E+02 | 7.23E+02 | 2.83E+01 | 4.39E+02 |
| Powell | Mean | 5.79E-02 | 2.43E+02 | 3.58E+02 | 2.25E+01 | 3.98E+02 | 2.98E+03 | 1.56E+04 | 1.49E+02 | 2.35E+03 |
| | Std | 4.07E-02 | 3.06E+02 | 5.01E+02 | 8.61E+00 | 9.86E+01 | 2.01E+03 | 3.57E+03 | 1.14E+02 | 1.10E+03 |
| | Min | 1.85E-02 | 1.67E+01 | 3.04E+01 | 8.69E+00 | 2.04E+02 | 1.76E+02 | 1.10E+04 | 1.99E+01 | 1.02E+03 |
| | Max | 1.76E-01 | 1.57E+03 | 2.49E+03 | 4.77E+01 | 6.25E+02 | 7.90E+03 | 2.35E+04 | 4.62E+02 | 6.55E+03 |

| | | PSOVA | PSO | BBPSO | BBPSOV | AGPSO | ELPSO | DNLPSO | MSPSO | GPSO |
|---|---|---|---|---|---|---|---|---|---|---|
| Dixon-Price | Mean | 4.77E+00 | 2.30E+06 | 2.63E+05 | 6.20E+04 | 2.68E+01 | 1.00E+01 | 1.68E+03 | 1.29E+05 | 2.32E+06 |
| | Std | 3.16E+00 | 7.71E+05 | 3.53E+05 | 6.68E+04 | 2.49E+01 | 0.00E+00 | 8.09E+03 | 3.74E+04 | 5.32E+05 |
| | Min | 6.70E-01 | 1.28E+06 | 7.86E+01 | 3.45E+03 | 3.92E+00 | 1.00E+01 | 6.82E-01 | 6.68E+04 | 1.57E+06 |
| | Max | 9.82E+00 | 5.13E+06 | 1.41E+06 | 3.45E+05 | 1.01E+02 | 1.00E+01 | 4.45E+04 | 2.01E+05 | 3.79E+06 |
| Sphere | Mean | 5.21E-08 | 1.47E+02 | 2.35E+01 | 1.77E+01 | 5.96E-03 | 5.12E+00 | 1.89E+00 | 4.58E+01 | 1.72E+02 |
| | Std | 1.03E-07 | 2.59E+01 | 2.57E+01 | 8.31E+00 | 5.45E-03 | 1.81E-15 | 6.02E+00 | 7.69E+00 | 2.10E+01 |
| | Min | 9.30E-10 | 8.46E+01 | 1.29E-02 | 5.80E+00 | 8.21E-04 | 5.12E+00 | 5.18E-07 | 3.11E+01 | 1.27E+02 |
| | Max | 3.22E-07 | 1.80E+02 | 1.05E+02 | 4.85E+01 | 2.36E-02 | 5.12E+00 | 2.43E+01 | 6.19E+01 | 2.19E+02 |
| Rotated Hyper-Ellipsoid | Mean | 1.88E-04 | 6.03E+05 | 1.21E+05 | 6.39E+04 | 3.18E+03 | 6.55E+01 | 1.02E+04 | 1.66E+05 | 6.46E+05 |
| | Std | 3.94E-04 | 1.29E+05 | 1.13E+05 | 3.74E+04 | 9.14E+03 | 4.34E-14 | 2.84E+04 | 3.10E+04 | 1.15E+05 |
| | Min | 3.84E-06 | 3.96E+05 | 4.46E+03 | 1.74E+04 | 1.81E+00 | 6.55E+01 | 1.63E-03 | 1.22E+05 | 4.34E+05 |
| | Max | 1.74E-03 | 8.79E+05 | 3.48E+05 | 1.78E+05 | 3.86E+04 | 6.55E+01 | 1.30E+05 | 2.28E+05 | 8.94E+05 |
| Sum Squares | Mean | 4.78E-06 | 1.48E+04 | 2.32E+03 | 1.61E+03 | 1.16E+02 | 1.00E+01 | 4.03E+02 | 3.72E+03 | 1.56E+04 |
| | Std | 8.15E-06 | 3.97E+03 | 1.53E+03 | 1.06E+03 | 2.19E+02 | 0.00E+00 | 8.88E+02 | 7.13E+02 | 2.30E+03 |
| | Min | 1.41E-07 | 7.65E+03 | 3.02E+02 | 4.81E+02 | 9.03E-02 | 1.00E+01 | 1.60E-03 | 2.46E+03 | 1.07E+04 |
| | Max | 3.11E-05 | 2.47E+04 | 5.71E+03 | 5.01E+03 | 9.00E+02 | 1.00E+01 | 3.51E+03 | 4.86E+03 | 1.96E+04 |
| Sum of Different Powers | Mean | 3.28E-18 | 3.43E-01 | 7.92E-10 | 4.45E-06 | 3.34E-14 | 3.42E-08 | 9.90E-08 | 8.49E-05 | 4.24E-01 |
| | Std | 1.01E-17 | 2.60E-01 | 3.68E-09 | 7.78E-06 | 9.51E-14 | 1.60E-07 | 3.18E-07 | 8.57E-05 | 3.32E-01 |
| | Min | 1.20E-22 | 2.36E-02 | 1.27E-15 | 6.77E-10 | 1.37E-18 | 9.17E-19 | 5.65E-37 | 4.08E-06 | 2.22E-02 |
| | Max | 5.33E-17 | 9.53E-01 | 2.01E-08 | 3.00E-05 | 3.84E-13 | 8.76E-07 | 1.58E-06 | 3.03E-04 | 1.28E+00 |
| Ackley | Mean | 3.27E+00 | 1.95E+01 | 1.85E+01 | 1.25E+01 | 5.84E+00 | 2.37E+00 | 6.89E+00 | 1.54E+01 | 1.99E+01 |
| | Std | 8.23E-01 | 5.71E-01 | 1.41E+00 | 1.95E+00 | 1.35E+00 | 4.82E-01 | 2.74E+00 | 6.32E-01 | 2.73E-01 |
| | Min | 2.08E+00 | 1.82E+01 | 1.42E+01 | 7.41E+00 | 3.46E+00 | 1.32E+00 | 2.81E+00 | 1.40E+01 | 1.93E+01 |
| | Max | 5.68E+00 | 2.05E+01 | 2.01E+01 | 1.55E+01 | 9.04E+00 | 3.36E+00 | 1.44E+01 | 1.67E+01 | 2.05E+01 |
| Griewank | Mean | 1.78E-02 | 5.28E+02 | 7.27E+01 | 5.71E+01 | 6.51E-01 | 2.11E+01 | 7.97E+00 | 1.51E+02 | 6.01E+02 |
| | Std | 2.43E-02 | 9.93E+01 | 8.93E+01 | 2.52E+01 | 2.93E-01 | 1.09E+02 | 2.38E+01 | 2.94E+01 | 6.90E+01 |
| | Min | 1.14E-06 | 3.09E+02 | 1.04E+00 | 1.90E+01 | 1.22E-01 | 7.88E-01 | 2.17E-02 | 9.05E+01 | 4.57E+02 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | **Max** | **8.50E-02** | 6.91E+02 | 3.62E+02 | 1.21E+02 | 1.08E+00 | 6.00E+02 | 1.30E+02 | 1.97E+02 | 7.76E+02 |
| **Powell** | **Mean** | **5.79E-02** | 1.40E+04 | 2.67E+03 | 7.78E+02 | 3.62E+01 | 5.00E+00 | 2.57E+01 | 1.51E+03 | 1.21E+04 |
| | **Std** | 4.07E-02 | 4.03E+03 | 1.97E+03 | 7.29E+02 | 5.09E+01 | **0.00E+00** | 5.14E+01 | 3.87E+02 | 2.71E+03 |
| | **Min** | **1.85E-02** | 6.44E+03 | 1.79E+02 | 9.35E+01 | 4.76E-01 | 5.00E+00 | 6.14E-02 | 7.92E+02 | 7.87E+03 |
| | **Max** | **1.76E-01** | 2.24E+04 | 6.93E+03 | 2.94E+03 | 1.79E+02 | 5.00E+00 | 2.07E+02 | 2.32E+03 | 1.86E+04 |

**Table 5.19: The Wilcoxon rank sum test results for all benchmark functions with *D* = 50**

| | **BA** | **DA** | **HS** | **FPA** | **MFO** | **ABC** | **CA** | **GA** |
|---|---|---|---|---|---|---|---|---|
| **Dixon** | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 |
| **Sphere** | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 |
| **Rotated Hyper-Ellipsoid** | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 |
| **Sum Squares** | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 |
| **Sum of Different Powers** | 3.02E-11 | 3.34E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 |
| **Ackley** | 3.02E-11 | 3.34E-11 | 1.78E-04 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.34E-11 | 3.02E-11 |
| **Griewank** | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 |
| **Powell** | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 |

| | **PSO** | **BBPSO** | **BBPSOV** | **AGPSO** | **ELPSO** | **DNLPSO** | **MSPSO** | **GPSO** |
|---|---|---|---|---|---|---|---|---|
| **Dixon** | 3.02E-11 | 3.02E-11 | 3.02E-11 | 2.15E-10 | 1.21E-12 | 1.11E-06 | 3.02E-11 | 3.02E-11 |
| **Sphere** | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 1.21E-12 | 3.02E-11 | 3.02E-11 | 3.02E-11 |
| **Rotated Hyper-Ellipsoid** | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 1.21E-12 | 3.02E-11 | 3.02E-11 | 3.02E-11 |
| **Sum Squares** | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 1.21E-12 | 3.02E-11 | 3.02E-11 | 3.02E-11 |
| **Sum of Different Powers** | 3.02E-11 | 3.34E-11 | 3.02E-11 | 3.02E-11 | 3.34E-11 | 2.87E-10 | 3.02E-11 | 3.02E-11 |
| **Ackley** | 3.02E-11 | 3.02E-11 | 2.15E-10 | **8.30E-01** | 3.52E-07 | **1.86E-01** | 3.02E-11 | 3.02E-11 |
| **Griewank** | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 1.21E-10 | 3.02E-11 | 3.02E-11 |
| **Powell** | 3.02E-11 | 3.02E-11 | 3.02E-11 | 3.02E-11 | 1.21E-12 | 3.02E-11 | 3.02E-11 | 3.02E-11 |

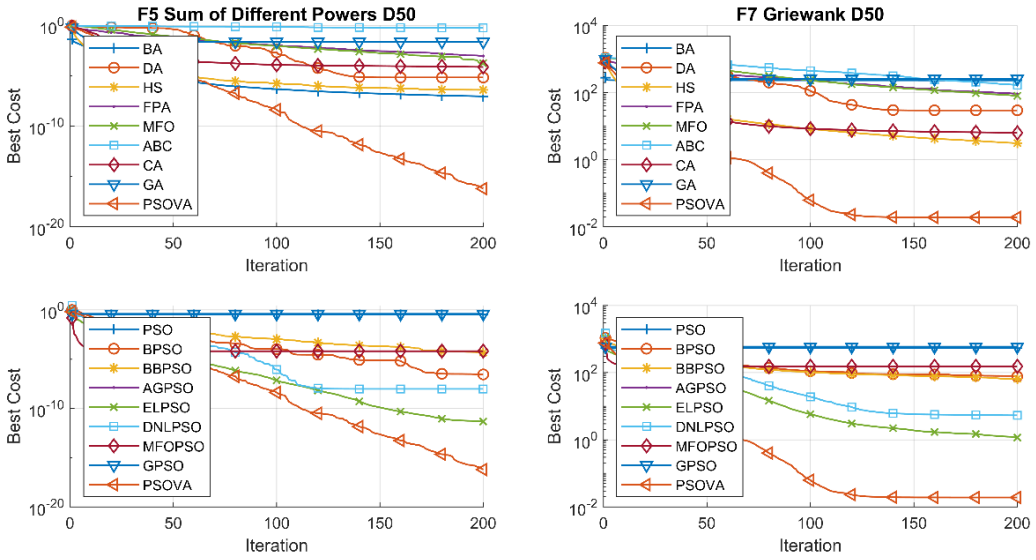



**Figure 5-3: Convergence curves for F5 and F7 in *D* = 50**

## 5.3 Chapter Summary

This study has employed the proposed PSOVA-based feature optimisation to address the classification of skin lesion problems. To explore the search space more broadly, several distinctive velocities updating strategies were proposed and integrated with the original PSO method. Two remote sub-swarm leaders are used to guide the exploration of the sub-swarms. Furthermore, to enhance diversifications, the study made use of probability

distributions and matrix representation. The premature convergence of the conventional PSO method can be improved by the proposed PSO algorithm. Furthermore, when it comes to addressing issues associated with discriminative feature selection, the PSOVA model is more efficient than other traditional search techniques and eight advanced PSO variations, as confirmed by the outcomes of the evaluation with several skin lesion and spam base datasets. Moreover, the proposed PSOVA model has also been proven more effective on the search mechanisms by statistical Wilcoxon rank sum test. There are several characteristics of the proposed PSO algorithm that can be modified in future studies. This includes its development with adaptive coefficients of $C1$ and $C2$ by using various nonlinear functions. This would allow the search process to focus on exploring the search space in early iterations before focusing on the local exploitation of the global best solution. Other medical datasets could also be used for the evaluation of the proposed PSO model, including breast and blood cancer, retinal disease, seizure detection and wine classification datasets. In addition, the proposed PSO model can also be used to identify optimal hyper-parameters of diverse classification models.

# CHAPTER 6
# PROPOSED METHODOLOGY 4 WITH ACPSO BASED FEATURE SELECTION

The above proposed PSO models could be further improved in a number of ways. Adaptive coefficients could be used to accelerate convergence. Therefore, another PSO model is proposed by addressing this factor. Specifically, fixed coefficients are used in the conventional PSO, whereas the proposed PSO model in this chapter employs adaptive search parameters to enable the search process to focus on global exploration from the beginning of the iterations and move on to local exploration in later iterations to accelerate convergence. The proposed model adaptively decreases *C1* and increases *C2* to increase local and global search capabilities. Specifically, *C1* and *C2* parameters are not fixed but dynamically adjusted at every iteration. We propose three different nonlinear strategies to generate adaptive decreasing and increasing coefficients and explore the search space more effectively.

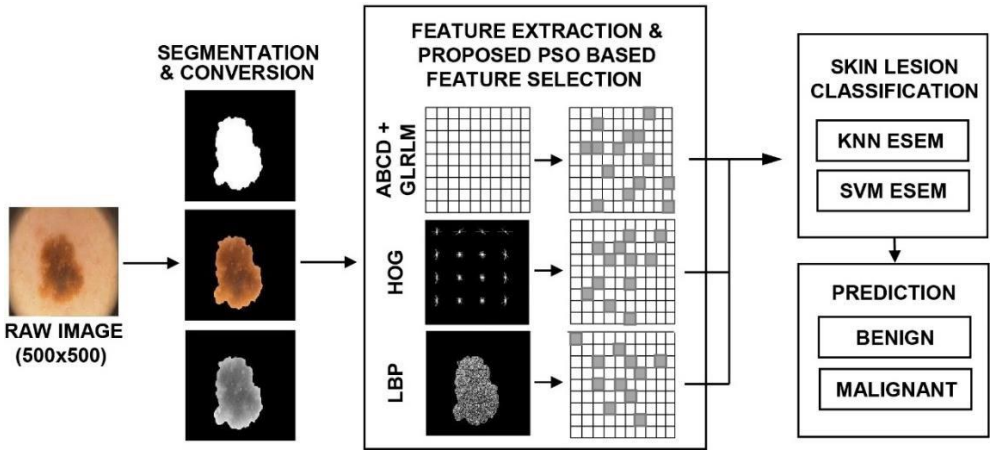
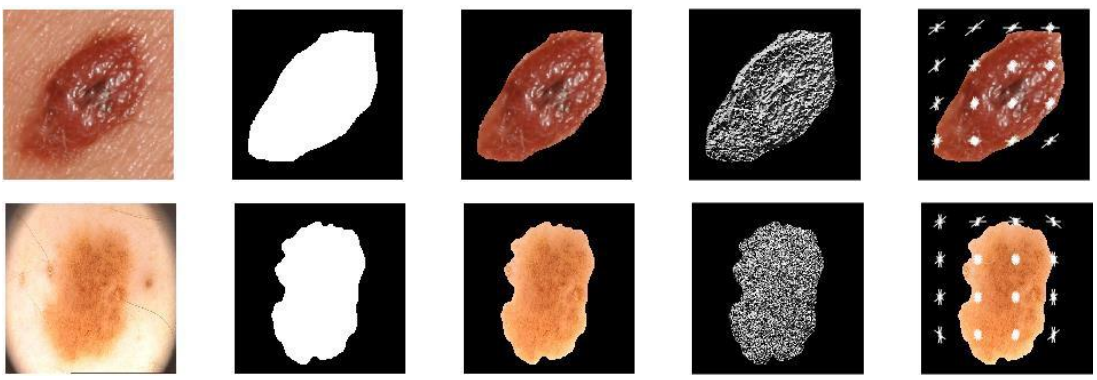



**Figure 6-1: The System Architecture for Enhanced Skin Lesion Classification**

## 6.1 Feature Extraction

Dermatologists have used ABCDE to help skin cancer diagnosis. In most cases, other related works have only employed ABCD rules of dermatology for feature extraction to make automatic skin cancer detection (Nachbar et al., 1994, Barata et al., 2015, Monisha et al., 2018, Navarro et al., 2018). In this research, multiple features have been extracted such as hand-crafted shape features, e.g. asymmetry, compactness, radial, border irregular, perimeter, solidity and extent; colour features, e.g. colour variance, entropy and skewness and high-level texture features. In this research, two low-level feature extraction methods have also been used for the representation of lesion texture, i.e. local binary patterns (LBP) and histogram of oriented gradients (HOG) operators. Specifically, this study uses an LBP operator to extract 236 features with a cell size of 191 × 191. This research makes use of the HOG operator as well, to extract more distinctive complementary texture features. The employed HOG operator is used to extract 324 features using the cell and block sizes of 110 × 110 and 4 × 4 respectively. An overlap of half, as well as two-thirds of the block size, is used in the experiments for the HOG operator. The HOG operator can achieve more efficient results with an overlap of half the block size than with an overlap of two-thirds of the block size. The outcomes of the experiments highlight that the HOG operator is highly efficient in the extraction of important contrast information (e.g. edges and corners). Therefore, these features are added in this chapter to improve the performance accuracy. These newly extracted LBP and HOG features are included on top of all previously extracted ABCD morphological, textural and colour feature sets, as shown in Table 4.1. The sample images illustrated below indicate the extracted LBP and HOG features.



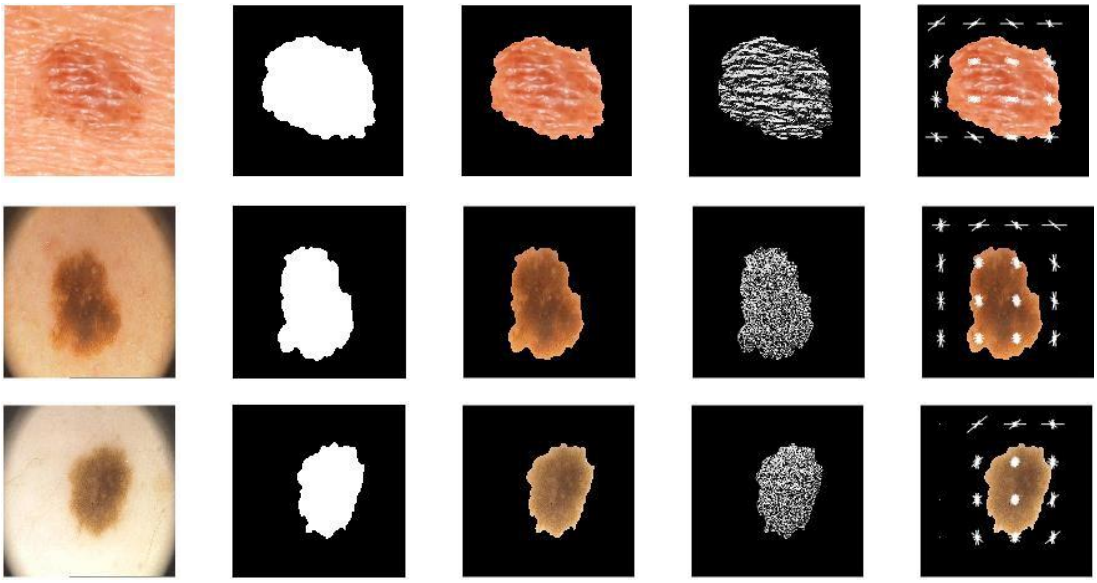


**Figure 6-2: Segmented lesion (third column), LBP lesion (fourth column) and HOG features (fifth column)**

## 6.2 The Proposed ACPSO Search Strategies for Feature Selection

Fundamentally, the workforce arrangement within a company inspires the idea of the proposed method. Each corporation is made up of a manager, a team leader and general workers; multiple workers are allocated to a task, since they do not have enough experience and employ different strategies; their skill, aptitude, working performance and quality are more likely to be imbalanced and progress with a range of diversity. In the beginning, each team leader is expected to have a similar experience to that of the manager. Therefore, the team leader oversees and guides the workers to make good progress. During the workers' progression, they intend to gain experience, and the possibility of being promoted. While the manager is less critical than the team leader and workers, the only goal of a manager is to guide the inexperienced group, to get them on the correct route with accurate information. Once the workers' task is completed, their responsibility is reduced; the manager then carries on the task for further inspection and submits it to the higher authorities. Although the workers are less important in the end, they will still be working as a group; each piece of information and idea can be useful for any further improvement. Therefore, they will keep collaborating and sharing ideas to complete the mission.

This research is motivated by the above cooperative strategies and proposes a PSO variant by utilising three different moving behaviours, including circle, helix and sine waveforms. They are used to lead each sub-swarm based search respectively. This algorithm is named as Adaptive Coefficient PSO (ACPSO). The purpose behind the three moving behaviours is to form different decreasing and increasing effects to update the coefficients, *C1* and *C2*. First a population of 50 particles is initialised, then all the particle is ranked based in their fitness scores. Three swarm leaders with similar fitness values but low correlation in positions are selected. Specifically, the best particle will be selected as $gBest$ and stored in the best memory. For the selection of the other two swarm leaders, a correlation coefficient is utilised; the two particles with competitive fitness which are the furthest and have low correlations with $gBest$ are selected. Furthermore, the population will be divided into three sub-swarms; each leader will then be assigned to lead each sub-swarm based search. The three adaptive exploration strategies, i.e. circle, waveform and helix functions, are then assigned into three sub-swarms respectively to lead its search process. Each sub-swarm based search uses the corresponding descending and ascending function to generate the $c1$ and $c2$ coefficients. Then the three sub-swarms are combined, and the new $gBest$ will be identified. Subsequently, a sub-dimension based exploration method will be conducted to further improve the $gBest$. Specifically, each particle in the swarm will be divided into three sub-dimensions which represent the upper, middle and bottom sections of skin lesions. The best solution of each sub-dimension is identified and then a new leader is generated by combining the three best leaders of the three sub-dimensions. This new global best solution is compared with the previous $gBest$. If the new leader has better fitness than that of the previous $gBest$, it will be used to replace $gBest$. Otherwise $gBest$ remains intact and the new leader is used to replace the worst particle in the swarm. Finally, the last second and third worst particles will be replaced by two newly randomly generated particles. These search processes terminate when the maximum iteration number is reached. The pseudo-code of ACPSO is shown in Algorithm 6.1.

**Algorithm 6.1: Pseudo-code of The Proposed ACPSO Model**

| | |
|---|---|
| **1** | **Start** |
| **2** | Initialize a population randomly (e.g. 50 particles); |
| **3** | Evaluate the population to identify the initial best leader**,** *gbest*; |
| **4** | |
| **5** | **While** (! Stop condition) do {// 125 iterations |
| **6** | Find second & third swarm leaders with comparable fitness but low correlation in position to $gBest$ using Eq. 5.1; |
| **7** | Divide the population into three sub-swarms, i.e. SS1, SS2 & SS3, with each sub-swarm led by one leader; |
| **8** | **While** (! Stop condition) do {// for three sub-swarm |
| **9** | **For** (each particle $x_i$ in sub-swarm 1) **do** { |
| **10** | Follow the sub-swarm leader using Eq. 4.1, 4.2 & 6.1, |
| **11** | Evaluate $x_i$ at the new location and update $pBest$; |
| **12** | Update the sub-swarm leader; |
| **13** | **} End For** |
| **14** | **For** (each particle $x_i$ in sub-swarm 2) **do** { |
| **15** | Follow the sub-swarm leader using Eq. 4.1, 4.2 & 6.2; |
| **16** | Evaluate $x_i$ at the new location and update $pBest$; |
| **17** | Update the sub-swarm leader; |
| **18** | **} End For** |
| **19** | **For** (each particle $x_i$ in sub-swarm 3) **do** { |
| **20** | Follow the sub-swarm leader using Eq. 4.1, 4.2 & 6.3-6.5; |
| **21** | Evaluate $x_i$ at the new location and update $pBest$; |
| **22** | Update the sub-swarm leader; |
| **23** | **} End For** |
| **24** | **} End While** |
| **25** | Combine all three sub-swarms and update $gBest$; |
| **26** | Divide each particle in the overall swarm into three sub-dimensions; |
| **27** | **While** (! Stop condition) do {// for three sub-dimensions |
| **28** | Follow the leader in each sub-dimension based search using Eq. 4.1-4.2.; |
| **29** | Identify each sub-dimension leader and combine them to generate a new swarm leader; |
| **30** | } **End While** |
| **31** | Combine three sub-dimensions and update the $gBest$ if the new leader is better; |
| **32** | Otherwise the new leader is used to replace the worst particle in the swarm; |
| **33** | Replace the last second and third worst particles with two newly generated particles; |
| **34** | **} End While** |
| **35** | Output $gBest$; |
| **36** | **End** |

### 6.2.1 Adaptive Coefficient Generation Using the Circle Function

The proposed algorithm has proposed three functions to generate adaptive coefficients *C1* and *C2*, including circle, sine and helix waveforms. Equation (6.1) defines the circle

function, which is used to generate decreasing *C1* and ascending *C2* in sub-swarm 1 with *C1* and *C2* ∈ [0, 2.5].

$$y = \sqrt{r^2 - x^2} \qquad x \in [0, 2.5] \tag{6.1}$$

In Equation 6.1, $x$ and $y$ represent the coordinates and $r$ denotes the radius of the circle. Based on Figure 6-3, the upper right quarter of the circle shape has been divided into $m$ portions to form the descending and ascending values for *C1* and *C2*, where $m$ represents the maximum number of iterations. For the generation of decreasing *C1*, the traverse starts from portion 1 (i.e. 90°) and ends in portion *m* (i.e. 0°). In each iteration, there is one value randomly selected in each portion. In this way, a series of descending *C1* is produced. Similarly, by traversing through from portion *m* (i.e. 0°) to portion 1 (i.e. 90°), a series of ascending *C2* is generated.

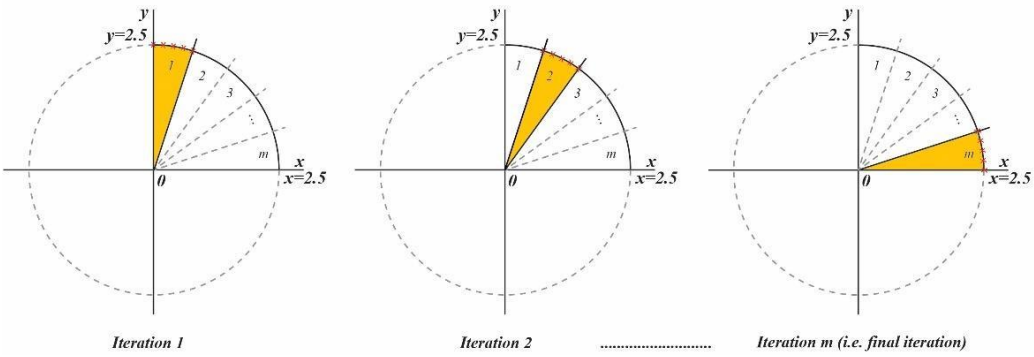



**Figure 6-3: The circle contour for the generation of descending *C1* and ascending *C2***

### 6.2.2 Adaptive Coefficient Generation Using the Sine Function

Similar ascending and descending processes can be generated using the sine function in sub-swarm 2. Equation (6.2) is used to form adaptive coefficient search on *C1* and *C2*.

$$y = 2.5 \times \sin(x) \qquad x \in [\frac{\pi}{2}, \pi] \tag{6.2}$$

Using Equation 6.2, a sine waveform is created as shown in Figure 6-4, where $x$ represents the input value range of $x \in [\frac{\pi}{2}, \pi]$. Similar to the coefficient generation using the circle function, the upper right quarter of the waveform has been divided into $m$ portions. In each iteration, a value is randomly selected in each portion. By going from

portion 1 (90°) to portion *m* (0°), a series of descending *C1* is generated, and vice versa from portion *m* (0°) to portion 1 (90°), a series of ascending *C2* is produced.

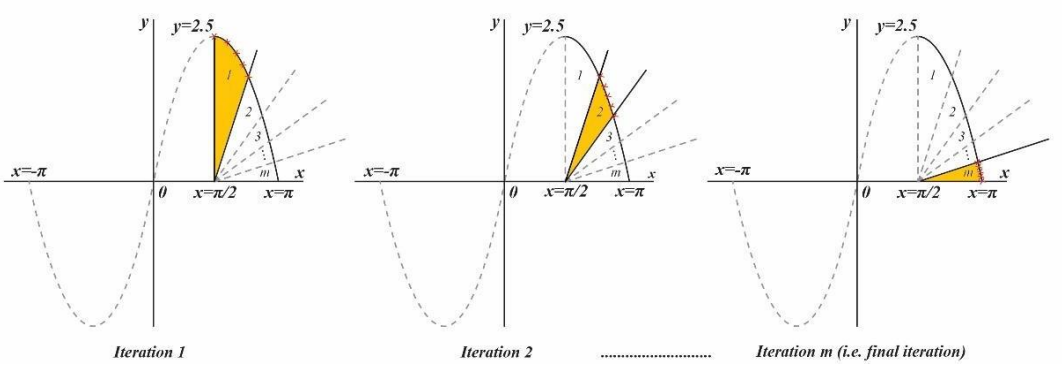



**Figure 6-4: The sine waveform for the generation of descending *C1* and ascending *C2***

### 6.2.3 Adaptive Coefficient Generation Using the Helix Function

In sub-swarm 3, the helix function defined in Equations (6.3) -(6.5) is used to generate the adaptive coefficients *C1* and *C2*. Nonetheless, the helix moving behaviours have been developed differently, since the shape of the helix is non-linear, and the circles are continually turning with a different angle, thus having multiple lines of a turning point. This also enables the model to use multiple layers as a more complex mechanism for the retrieval of both coefficients. The top half of the helix waveform is used for the adaptive coefficient generation as shown in Figure 6.5.

$$t \in [0, 389.5575]$$

$$x = 0.0065 \times t \times \cos(2 \times t) \qquad x \in [-2.522, 2.532] \tag{6.3}$$

$$y = 0.0065 \times t \times \sin(2 \times t) \qquad y \in [0, 2.517] \tag{6.4}$$

$$z = t \tag{6.5}$$

The helix function defined in the Equations (6.3)-(6.5) is used to generates $m$ layers as shown in Figure 6-5 where $t \in [0, 389.5575]$. First, both top left and right sides of the waveform are divided into *m* portions respectively. For the generation of descending *C1*, a random value is selected from portion 1 from both right and left sides in layer 1 for iteration 1. For iteration 2, a random value is selected from portion 2 from both right and left sides in layer 2. Eventually for the last iteration *m*, a random value is selected from

portion index *m* in layer *m*. In this way, a series of descending *C1* is generated. A reverse process is use for the generation of ascending coefficient *C2*. Specifically, for the generation of increasing *C2*, a random value is selected from portion *m* in both right and left sides in layer *m* in the first iteration. Then similarly, in the second iteration, a random value is selected from portion *m*-1 in both right and left sides in layer *m*-1. For the last iteration, a random value is selected from portion 1 in both right and left sides in layer 1. This process ensures the generation of a series of ascending *C2*. Therefore, the entire search process guided by $pBest$ and $gBest$ will be adjusted, i.e. in early iterations, the search focuses more on global exploration and in the final iterations, the search process concentrates more on local exploitation.

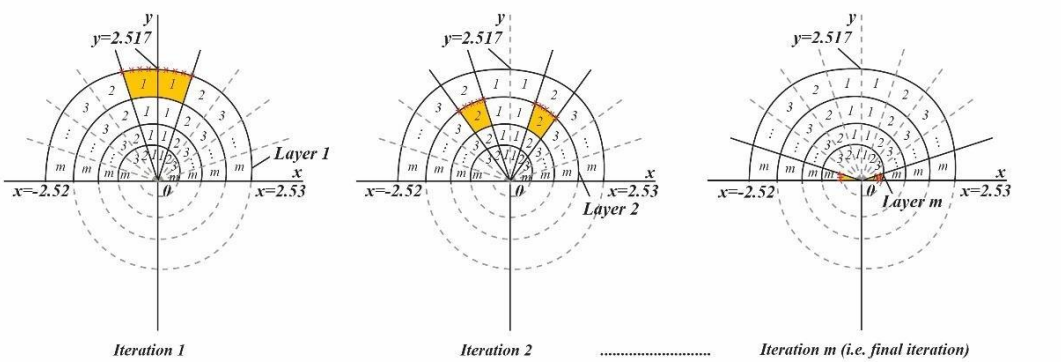



**Figure 6-5: The helix waveform for the generation of descending *C1* and ascending *C2***

The three functions listed above enable the generation of adaptive descending *C1* and ascending *C2*. They are used to guide each sub-swarm-based search. Each particle in the sub-swarm will be evaluated using the fitness evaluation defined in Equation (4.2).

## 6.3 Evaluation

In this experiment, several baseline methods as those mentioned in Section 4.3, and several PSO variants as those used in Section 5.2 have been nominated for comparison. Several additional PSO variants, including FBPSO, FSBPSO and MPSO, are also employed for comparison. The original 146 features, consisting of the hand-crafted high-level features mentioned in Section 4.1, and the additional LBP and HOG features, are

used to test each feature selection method. Each optimisation method will be evaluated using single and ensemble classification techniques for skin lesion classification.

This research utilised two dermoscopy image databases, i.e. the Edinburgh Research and Innovation (Dermofit) database and Dermatology Service of Hospital Pedro Hispano (PH2). The Dermofit Image Library consists of a total of 1,300 skin lesion images with ten skin lesion types, including e.g. Actinic Keratosis, Melanocytic Nevus, Malignant Melanomas, Seborrheic Keratosis and Basal Cell Carcinoma. Some of these ten lesion types can also be treated as two classes, i.e. benign and cancerous. Therefore, 190 benign and 174 cancerous images have been selected from this database for evaluation. As mentioned earlier, the PH2 database consists of a total of 200 skin lesion images with three skin lesion types including 80 common nevi (benign), 80 atypical nevi, and 40 melanoma (cancer) cases. Therefore, 80 benign and 40 cancerous images have been selected, which are combined with images from Dermofit for evaluation of the ACPSO. Therefore, a total of 270 benign (190 Dermofit and 80 PH2) and 214 cancerous (174 Dermofit and 40 PH2) images are selected from the two databases for our experiments. To compare with related works, another experiment using the PH2 dataset is also conducted. Both experiments were conducted with a ratio of 80:20 for the training and testing split, i.e. 387 and 97 images have been utilised for the above combined dataset; and 160 and 40 images for a PH2 dataset for training and testing, respectively.

**Table 6.1: The combined dataset for skin cancer classification for this study**

| Dataset | Benign | Cancer |
|---|---|---|
| **PH2** | 80 | 40 |
| **Dermofit** | 190 | 174 |

Two classifiers, i.e. Ensemble K-Nearest Neighbour (ESEM-KNN) and Ensemble Support Vector Machine (ESEM-SVM), are employed for verification of the outcomes. All the methods are based on the same experimental setting, i.e. 500 × 500 image size, population size of 50, dimension of 146 with ABCD+GLRLM, 236 of LBP or 324 HOG features, and 500 iterations with 30 runs. Furthermore, the maximum number of fitness evaluation of 50 populations × 500 iterations, is employed for each method to ensure the comparison is undertaken on equality. The proposed ACPSO model employs a much-

reduced number of iterations, i.e. 50 population × 125 iterations × 4 function evaluations = 25,000 function evaluations, owing to the extra fitness evaluation conducted for a local search of lesion features to ensure fair comparison. The experimental details are provided below.

### 6.3.1 Parameter Settings

In this section, we examine parameter settings of ACPSO and other methods. Table 6.2 shows the detailed parameter settings identified for ACPSO and other newly employed optimisation methods. The ACPSO model follows the conventional PSO model parameter settings, with additional constricted parameter adjustments made based on a trial-and-error basis. The ACPSO model utilises an inertia weight of 0.65 to identify the effect of the prior speed on repetitive searches. As those in the initial PSO algorithm, initial acceleration constants $C1$ = 2.5 and $C2$ = 0.5 are employed to expedite convergence. For other search methods, the parameter settings are employed as stipulated in the previous investigations.

**Table 6.2: Parameter settings of the proposed and other algorithms**

| Algorithms | Parameters |
|---|---|
| **F-BPSO** | Inertia weight=0.689343, *C1*=*C2*=1.42694, initial mutation rate=0.11, final mutation rate=0.01 |
| **FS-BSPO** | Inertia weight=0.689343, *C1*=C2=1.42694, initial mutation rate=0.11, final mutation rate=0.01 |
| **MPSO** | Initial Inertia weight=0.4, final inertia weight=0.9, *C1*=final *C2*=2.5, initial *C2*=final *C1*=0.5 |
| **TAC-PSO-MS** | Standard deviation of Gaussian distribution=1.3, scale parameter of Cauchy distribution=2, inertia weight=0.9, initial *C1*=final *C2*=2.5, initial *C2*=final *C1*=1.5 |
| **ASCA-PSO** | Maximum velocity=1.0, Minimum velocity=-1.0, acceleration constants *C1*=C2=0.5, constant a=2 |
| **ACPSO** | inertia weight=0.65, initial *C1*=2.5, initial *C2*= 0.5 |

### 6.3.2 Evaluation Using the Combined Dataset

This experiment employ both above-mentioned combined dataset and the PH2 dataset for the assessment of each optimization method. A total of 30 runs are conducted for each model. The evaluation results for the combined dataset are shown In Table 6.4 In combination with KNN-based ensemble and SVM-based ensemble classifiers under both 10-fold and hold-out validations, the proposed ACPSO model attains higher average GM

results than those of all other algorithms. Indeed, the outcome of the KNN-based ensemble with 10-fold classification suggests that ACPSO outperforms BA, CA, CS, DA, HS, ABC, FPA, MFO, PSO, GPSO, MPSO, GMPSO, FB-PSO, ELPSO, MFOPSO, AGPSO, FS-BPSO, DNLPSO and BBPSOV, respectively. Furthermore, for KNN-based ensemble with hold-out validation, it also outperforms BA, CA, CS, DA, HS, ABC, FPA, MFO, PSO, GPSO, MPSO, GMPSO, FB-PSO, ELPSO, MFOPSO, AGPSO, FS-BPSO, DNLPSO and BBPSOV, respectively. When the SVM-based ensemble classifier combined with 10-fold validation was applied, the proposed algorithm outperforms BA, CA, CS, DA, HS, ABC, FPA, MFO, PSO, GPSO, MPSO, GMPSO, FB-PSO, ELPSO, MFOPSO, AGPSO, FS-BPSO, DNLPSO and BBPSOV, respectively. And lastly, for the SVM-based ensemble model integrated with the hold-out validation, ACPSO outperforms BA, CA, CS, DA, HS, ABC, FPA, MFO, PSO, GPSO, MPSO, GMPSO, FB-PSO, ELPSO, MFOPSO, AGPSO, FS-BPSO, DNLPSO and BBPSOV, respectively. The statistical results shown in Table 6.4 indicate that the proposed model is statistically significantly better than other methods for nearly all the test cases, except for ThBPSO which shows similar mean results to those of the proposed model when combined with SVM-based ensemble under 10-fold validation.

**Table 6.3: Average performance correlation of each algorithm over 30 runs for the combined dataset**

| | **ESEM KNN** | | **ESEM SVM** | |
|---|---|---|---|---|
| | **10-fold** | **Hold-out** | **10-fold** | **Hold-out** |
| **BA** | 1.98% | 2.13% | 1.30% | 1.40% |
| **CA** | 2.52% | 2.65% | 1.13% | 1.24% |
| **CS** | 2.30% | 2.72% | 0.83% | 0.96% |
| **DA** | 1.56% | 2.09% | 0.92% | 1.08% |
| **HS** | 1.91% | 2.27% | 0.86% | 1.28% |
| **ABC** | 1.79% | 1.74% | 0.65% | 0.84% |
| **FPA** | 2.45% | 2.53% | 1.23% | 1.44% |
| **MFO** | 1.90% | 2.22% | 0.89% | 1.24% |
| **PSO** | 2.24% | 2.33% | 1.10% | 1.18% |
| **BBPSO** | 2.07% | 2.18% | 0.91% | 1.01% |
| **ThBPSO** | 1.98% | 2.19% | 0.50% | 0.80% |
| **GPSO** | 2.00% | 2.40% | 0.73% | 0.88% |
| **MPSO** | 2.38% | 2.52% | 0.89% | 0.93% |
| **GMPSO** | 2.41% | 2.58% | 0.84% | 1.16% |
| **FB-PSO** | 1.21% | 1.83% | 0.63% | 0.75% |
| **ELPSO** | 1.47% | 1.85% | 0.64% | 0.92% |
| **MFOPSO** | 2.10% | 2.30% | 1.26% | 1.53% |
| **AGPSO** | 2.20% | 2.32% | 0.95% | 1.17% |
| **FS-BPSO** | 1.60% | 2.00% | 0.62% | 1.03% |
| **DNLPSO** | 2.08% | 2.62% | 1.25% | 1.24% |
| **BBPSOV** | 2.44% | 2.85% | 0.71% | 0.99% |

**Table 6.4: Average performance of each algorithm over 30 runs for the combined dataset**

| | ESEM KNN | | ESEM SVM | |
|---|---|---|---|---|
| | **10-fold** | **Hold-out** | **10-fold** | **Hold-out** |
| **BA** | 0.9362 | 0.9328 | 0.9825 | 0.9809 |
| **CA** | 0.9308 | 0.9276 | 0.9842 | 0.9825 |
| **CS** | 0.9330 | 0.9269 | 0.9872 | 0.9853 |
| **DA** | 0.9404 | 0.9332 | 0.9863 | 0.9841 |
| **HS** | 0.9369 | 0.9314 | 0.9869 | 0.9821 |
| **ABC** | 0.9381 | 0.9367 | 0.9890 | 0.9865 |
| **FPA** | 0.9315 | 0.9288 | 0.9832 | 0.9805 |
| **MFO** | 0.9370 | 0.9319 | 0.9866 | 0.9825 |
| **PSO** | 0.9336 | 0.9308 | 0.9845 | 0.9831 |
| **BBPSO** | 0.9353 | 0.9323 | 0.9864 | 0.9848 |
| **ThBPSO** | 0.9362 | 0.9322 | 0.9905 | 0.9869 |
| **GPSO** | 0.9360 | 0.9301 | 0.9882 | 0.9861 |
| **MPSO** | 0.9322 | 0.9289 | 0.9866 | 0.9856 |
| **GMPSO** | 0.9319 | 0.9283 | 0.9871 | 0.9833 |
| **FB-PSO** | 0.9439 | 0.9358 | 0.9892 | 0.9874 |
| **ELPSO** | 0.9413 | 0.9356 | 0.9891 | 0.9857 |
| **MFOPSO** | 0.9350 | 0.9311 | 0.9829 | 0.9796 |
| **AGPSO** | 0.9340 | 0.9309 | 0.9860 | 0.9832 |
| **FS-BPSO** | 0.9400 | 0.9341 | 0.9893 | 0.9846 |
| **DNLPSO** | 0.9352 | 0.9279 | 0.9830 | 0.9825 |
| **BBPSOV** | 0.9316 | 0.9256 | 0.9884 | 0.9850 |
| **ACPSO** | **0.9560** | **0.9541** | **0.9955** | **0.9949** |

**Table 6.5: The *p-values* of the Wilcoxon rank sum test using the combined dataset**

| | **BA** | **CA** | **CS** | **DA** | **HS** | **ABC** | **FPA** | **MFO** | **PSO** | **BBPSO** | **ThBPSO** |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **ESEM KNN Wilcoxon 10F** | 7.97E-06 | 1.15E-07 | 1.10E-08 | 4.33E-05 | 3.06E-08 | 1.42E-05 | 1.65E-07 | 8.99E-08 | 1.20E-07 | 4.41E-07 | 2.00E-08 |
| **ESEM KNN Wilcoxon HO** | 4.03E-07 | 2.38E-09 | 5.56E-10 | 7.50E-11 | 7.37E-10 | 4.80E-07 | 1.75E-09 | 1.58E-07 | 1.26E-08 | 2.12E-07 | 3.39E-09 |
| **ESEM SVM Wilcoxon 10F** | 1.64E-05 | 1.48E-05 | 7.70E-04 | 2.42E-07 | 3.74E-04 | 3.45E-03 | 1.09E-05 | 9.79E-07 | 3.90E-05 | 1.36E-04 | **5.38E-02** |
| **ESEM SVM Wilcoxon HO** | 2.96E-08 | 8.35E-08 | 7.64E-06 | 1.43E-10 | 1.33E-07 | 6.88E-06 | 2.21E-08 | 5.92E-09 | 9.25E-06 | 6.42E-06 | 1.90E-04 |

| | **GPSO** | **MPSO** | **GMPSO** | **FB-PSO** | **ELPSO** | **MFOPSO** | **AGPSO** | **FS-BPSO** | **DNLPSO** | **BBPSOV** |
|---|---|---|---|---|---|---|---|---|---|---|
| **ESEM KNN Wilcoxon 10F** | 3.77E-07 | 8.25E-08 | 8.82E-10 | 1.23E-03 | 2.19E-05 | 8.99E-08 | 8.25E-08 | 2.77E-07 | 1.42E-08 | 1.61E-09 |
| **ESEM KNN Wilcoxon HO** | 1.90E-08 | 4.38E-09 | 3.83E-10 | 3.91E-09 | 1.17E-08 | 1.58E-07 | 4.38E-09 | 1.78E-07 | 2.67E-10 | 8.65E-11 |
| **ESEM SVM Wilcoxon 10F** | 3.53E-03 | 1.21E-05 | 1.85E-05 | 2.57E-03 | 2.96E-05 | 9.79E-07 | 1.21E-05 | 2.31E-02 | 4.78E-06 | 8.55E-04 |
| **ESEM SVM Wilcoxon HO** | 3.13E-05 | 1.00E-07 | 1.00E-07 | 4.73E-04 | 9.94E-09 | 5.92E-09 | 1.00E-07 | 8.48E-05 | 2.00E-07 | 2.82E-06 |

**Table 6.6: The training computational cost and the average number of features selected using the combined dataset**

| | **ACPSO** | **BA** | **CA** | **CS** | **DA** | **HS** | **ABC** | **FPA** | **MFO** | **PSO** | **BBPSO** |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Average no. of selected features** | 62.65 | 74.35 | 67.38 | 75.23 | 78.09 | 71.86 | 70.78 | 70.73 | 65.93 | 73.58 | 73.17 |
| **Training cost (seconds)** | 3798.22 | 3816.57 | 3929.63 | 3927.12 | 3813.44 | 3931.89 | 3906.70 | 3806.12 | 3905.52 | 3878.16 | 3937.50 |

| | GPSO | MPSO | GMPSO | FB-PSO | ELPSO | MFOPSO | AGPSO | FS-BPSO | DNLPSO | ThBPSO | BBPSOV |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Average no. of selected features** | 72.37 | 59.62 | 74.26 | 70.29 | 77.33 | 78.54 | 65.31 | 75.60 | 69.73 | 53.13 | 64.67 |
| **Training cost (seconds)** | 3945.48 | 3998.31 | 3878.48 | 3872.37 | 3970.20 | 3932.88 | 3879.13 | 4011.23 | 3860.92 | 4048.19 | 3981.04 |

### 6.3.3 Evaluation Using the PH2 Dataset

The PH2 dataset is also used for evaluating the proposed model. Again, the performances are categorised using the average of over 30 runs for each method for comparison. The KNN-based and SVM-based ensemble classifiers with both 10-fold and hold-out validations are used to test the feature subsets retrieved by each algorithm. The detailed results are shown in Table 6.8, with the optimum results for individual rows being emphasised in bold. In comparison with all the other algorithms, the proposed ACPSO model achieves the highest average GM rates when combined with both KNN-based and SVM-based ensembles using 10-fold and hold-out validations. Based on the results obtained using KNN-based ensemble with 10-fold validation, ACPSO reliably outperforms 20 competing algorithms i.e. BA, CA, CS, DA, HS, ABC, FPA, MFO, PSO, GPSO, MPSO, GMPSO, FB-PSO, ELPSO, MFOPSO, AGPSO, FS-BPSO, DNLPSO and BBPSOV, respectively. Furthermore, for the KNN-based ensemble with the hold-out validation, ACPSO also outperforms BA, CA, CS, DA, HS, ABC, FPA, MFO, PSO, GPSO, MPSO, GMPSO, FB-PSO, ELPSO, MFOPSO, AGPSO, FS-BPSO, DNLPSO and BBPSOV, respectively. When the SVM-based ensemble classifier combined with 10-fold validation is applied, the proposed algorithm outperforms BA, CA, CS, DA, HS, ABC, FPA, MFO, PSO, GPSO, MPSO, GMPSO, FB-PSO, ELPSO, MFOPSO, AGPSO, FS-BPSO, DNLPSO and BBPSOV, respectively. And lastly, for the SVM-based ensemble classifier with hold-out validation, ACPSO outperforms BA, CA, CS, DA, HS, ABC, FPA, MFO, PSO, GPSO, MPSO, GMPSO, FB-PSO, ELPSO, MFOPSO, AGPSO, FS-BPSO, DNLPSO and BBPSOV, respectively. As shown in Table 6.6, among the baseline methods, FS-BPSO obtains impressive performances. ACPSO outperforms other methods statistically in most of the test cases.

**Table 6.7: Average performance correlation of each algorithm over 30 runs for the PH2 dataset**

| | ESEM KNN | | ESEM SVM | |
|---|---|---|---|---|
| | **10-fold** | **Hold-out** | **10-fold** | **Hold-out** |
| **BA** | 3.19% | 2.04% | 3.02% | 3.76% |
| **CA** | 3.66% | 1.71% | 1.49% | 3.02% |
| **CS** | 3.32% | 2.10% | 1.05% | 2.11% |
| **DA** | 3.63% | 1.65% | 1.74% | 3.22% |
| **HS** | 2.23% | 0.92% | 1.54% | 2.78% |
| **ABC** | 3.41% | 1.98% | 1.39% | 2.45% |
| **FPA** | 3.93% | 2.49% | 2.64% | 3.72% |
| **MFO** | 2.76% | 1.88% | 1.10% | 1.99% |
| **PSO** | 3.29% | 2.17% | 1.49% | 2.59% |
| **BBPSO** | 3.61% | 2.50% | 1.56% | 2.84% |
| **ThBPSO** | 2.39% | 1.20% | 1.75% | 2.26% |
| **GPSO** | 2.70% | 1.42% | 1.33% | 2.85% |
| **MPSO** | 3.12% | 1.82% | 2.38% | 2.53% |
| **GMPSO** | 3.49% | 2.11% | 1.90% | 3.19% |
| **FB-PSO** | 3.36% | 1.31% | 1.51% | 2.96% |
| **ELPSO** | 3.41% | 1.79% | 1.52% | 2.99% |
| **MFOPSO** | 3.59% | 2.57% | 1.86% | 2.03% |
| **AGPSO** | 3.71% | 2.45% | 1.21% | 2.07% |
| **FS-BPSO** | 1.96% | 0.61% | 1.21% | 2.50% |
| **DNLPSO** | 2.55% | 1.83% | 1.63% | 2.63% |
| **BBPSOV** | 3.54% | 1.99% | 1.25% | 2.31% |

**Table 6.8: Average performance of each algorithm over 30 runs for the PH2 dataset**

| | ESEM KNN | | ESEM SVM | |
|---|---|---|---|---|
| | **10-fold** | **Hold-out** | **10-fold** | **Hold-out** |
| **BA** | 0.9407 | 0.9534 | 0.9465 | 0.9355 |
| **CA** | 0.9360 | 0.9567 | 0.9618 | 0.9429 |
| **CS** | 0.9394 | 0.9528 | 0.9662 | 0.9520 |
| **DA** | 0.9363 | 0.9573 | 0.9593 | 0.9409 |
| **HS** | 0.9503 | 0.9646 | 0.9613 | 0.9453 |
| **ABC** | 0.9385 | 0.9540 | 0.9628 | 0.9486 |
| **FPA** | 0.9333 | 0.9489 | 0.9503 | 0.9359 |
| **MFO** | 0.9450 | 0.9550 | 0.9657 | 0.9532 |
| **PSO** | 0.9397 | 0.9521 | 0.9618 | 0.9472 |
| **BBPSO** | 0.9365 | 0.9488 | 0.9611 | 0.9447 |
| **ThBPSO** | 0.9487 | 0.9618 | 0.9592 | 0.9505 |
| **GPSO** | 0.9456 | 0.9596 | 0.9634 | 0.9446 |
| **MPSO** | 0.9414 | 0.9556 | 0.9529 | 0.9478 |
| **GMPSO** | 0.9377 | 0.9527 | 0.9577 | 0.9412 |
| **F-BPSO** | 0.9390 | 0.9607 | 0.9616 | 0.9435 |
| **ELPSO** | 0.9385 | 0.9559 | 0.9615 | 0.9432 |
| **MFOPSO** | 0.9367 | 0.9481 | 0.9581 | 0.9528 |
| **AGPSO** | 0.9355 | 0.9493 | 0.9646 | 0.9524 |
| **FS-BPSO** | 0.9530 | 0.9677 | 0.9646 | 0.9481 |
| **DNLPSO** | 0.9471 | 0.9555 | 0.9604 | 0.9468 |
| **BBPSOV** | 0.9372 | 0.9539 | 0.9642 | 0.9500 |
| **ACPSO** | **0.9726** | **0.9738** | **0.9767** | **0.9731** |

**Table 6.9: The *p*-values of the Wilcoxon rank sum test using PH2 dataset**

| | BA | CA | CS | DA | HS | ABC | FPA | MFO | PSO | BBPSO | ThBPSO |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **ESEM KNN Wilcoxon 10F** | 6.86E-08 | 4.37E-08 | 2.57E-07 | 4.85E-08 | 2.07E-05 | 2.17E-07 | 1.24E-07 | 5.89E-08 | 3.51E-07 | 8.95E-07 | 1.70E-05 |
| **ESEM KNN Wilcoxon HO** | 2.29E-10 | 2.56E-05 | 3.80E-07 | 3.11E-07 | 8.05E-04 | 6.92E-06 | 2.82E-09 | 1.54E-08 | 3.90E-07 | 2.62E-09 | 8.68E-04 |
| **ESEM SVM Wilcoxon 10F** | 1.46E-05 | 1.37E-02 | **9.58E-02** | 6.69E-04 | 5.28E-03 | 5.36E-03 | 4.96E-04 | 3.96E-04 | 4.69E-03 | 4.55E-03 | 1.15E-03 |
| **ESEM SVM Wilcoxon HO** | 3.38E-08 | 1.03E-07 | 5.69E-05 | 2.05E-08 | 1.97E-07 | 7.19E-06 | 5.29E-09 | 3.26E-05 | 5.49E-06 | 2.67E-06 | 1.93E-05 |

| | GPSO | MPSO | GMPSO | F-BPSO | ELPSO | MFOPSO | AGPSO | FS-BPSO | DNLPSO | BBPSOV |
|---|---|---|---|---|---|---|---|---|---|---|
| **ESEM KNN Wilcoxon 10F** | 8.75E-07 | 1.87E-07 | 7.25E-09 | 2.29E-05 | 2.56E-07 | 5.89E-08 | 1.87E-07 | **5.04E-02** | 7.67E-07 | 4.16E-08 |
| **ESEM KNN Wilcoxon HO** | 2.91E-04 | 5.24E-11 | 1.39E-08 | 3.52E-03 | 1.44E-09 | 1.54E-08 | 5.24E-11 | **6.69E-01** | 1.60E-06 | 1.85E-09 |
| **ESEM SVM Wilcoxon 10F** | 9.51E-03 | **8.91E-02** | 2.42E-03 | 1.38E-02 | 1.86E-02 | 3.96E-04 | **8.91E-02** | **9.72E-02** | 3.86E-04 | **7.43E-02** |
| **ESEM SVM Wilcoxon HO** | 1.40E-07 | 2.36E-05 | 1.64E-07 | 1.17E-06 | 1.36E-07 | 3.26E-05 | 2.36E-05 | 1.34E-05 | 9.71E-07 | 1.84E-05 |

**Table 6.10: The training computational cost and the average number of features selected using PH2 dataset**

| | ACPSO | BA | CA | CS | DA | HS | ABC | FPA | MFO | PSO | BBPSO |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Average no. of selected features** | 65.04 | 81.57 | 75.08 | 78.41 | 83.13 | 73.89 | 78.12 | 75.45 | 69.65 | 83.61 | 82.01 |
| **Training cost (seconds)** | 3799.03 | 3823.17 | 3933.10 | 3937.52 | 3817.28 | 3933.13 | 3914.06 | 3811.73 | 3915.43 | 3885.54 | 3942.80 |

| | GPSO | MPSO | GMPSO | FB-PSO | ELPSO | MFOPSO | AGPSO | FS-BPSO | DNLPSO | ThBPSO | BBPSOV |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Average no. of selected features** | 73.46 | 66.76 | 80.18 | 70.68 | 85.61 | 84.20 | 71.35 | 78.71 | 71.01 | 54.38 | 74.79 |
| **Training cost (seconds)** | 3950.32 | 4005.86 | 3881.97 | 3876.67 | 3976.87 | 3935.33 | 3881.85 | 4020.29 | 3863.69 | 4054.51 | 3981.54 |

### 6.3.4 Evaluation Using the UCI Breast Cancer Dataset

To determine the impact of the proposed PSO algorithm on feature optimisation, an extra dataset from a different medical domain has also been employed for comparison, i.e. the UCI breast cancer dataset comprising 569 samples, 30 features and two categories (positive and negative). An aspect ratio of 80:20 is used as the training and testing split. In detail, the training and testing sets consist of 455 and 114 samples respectively. For each method, 30 runs were carried out, with the experimental set-up consisting of 30 dimensions, 50 populations and 500 repetitions. The number of function assessments of 50 population × 500 maximum number of repetitions is used as the stopping criteria. Table 6.11 indicates the average GM scores across 30 runs for each method, while Table 6.12 lists the outcomes of the Wilcoxon rank sum test. As indicated in Tables 6.7-6.8, the proposed ACPSO model achieves the best performances and outperforms all other methods for all the test cases statistically. Overall, when it comes to the formulation of

solutions associated with discriminative feature selection in relation to the breast cancer dataset, the statistical results confirm that the proposed PSO model is more efficient than alternative approaches.

**Table 6.11: Average performance of each algorithm over 30 runs for the breast cancer dataset**

| | ESEM KNN | | ESEM SVM | |
|---|---|---|---|---|
| | **10-fold** | **Hold-out** | **10-fold** | **Hold-out** |
| **BA** | 0.9583 | 0.9620 | 0.9769 | 0.9793 |
| **CA** | 0.9588 | 0.9629 | 0.9767 | 0.9795 |
| **CS** | 0.9583 | 0.9623 | 0.9784 | 0.9808 |
| **DA** | 0.9598 | 0.9634 | 0.9794 | 0.9818 |
| **HS** | 0.9592 | 0.9635 | 0.9841 | 0.9805 |
| **ABC** | 0.9597 | 0.9633 | 0.9785 | 0.9808 |
| **FPA** | 0.9606 | 0.9642 | 0.9785 | 0.9805 |
| **MFO** | 0.9582 | 0.9622 | 0.9795 | 0.9816 |
| **PSO** | 0.9600 | 0.9647 | 0.9795 | 0.9816 |
| **BBPSO** | 0.9582 | 0.9622 | 0.9795 | 0.9816 |
| **ThBPSO** | 0.9603 | 0.9635 | 0.9810 | 0.9833 |
| **GPSO** | 0.9592 | 0.9630 | 0.9796 | 0.9818 |
| **MPSO** | 0.9606 | 0.9639 | 0.9776 | 0.9798 |
| **GMPSO** | 0.9592 | 0.9630 | 0.9811 | 0.9831 |
| **F-BPSO** | 0.9620 | 0.9615 | 0.9789 | 0.9815 |
| **ELPSO** | 0.9565 | 0.9635 | 0.9766 | 0.9825 |
| **MFOPSO** | 0.9601 | 0.9634 | 0.9778 | 0.9800 |
| **AGPSO** | 0.9584 | 0.9628 | 0.9802 | 0.9823 |
| **FS-BPSO** | 0.9605 | 0.9610 | 0.9775 | 0.9785 |
| **DNLPSO** | 0.9440 | 0.9484 | 0.9760 | 0.9780 |
| **BBPSOV** | 0.9591 | 0.9628 | 0.9784 | 0.9805 |
| **ACPSO** | **0.9700** | **0.9688** | **0.9864** | **0.9856** |

**Table 6.12: The *p-values* of the Wilcoxon rank sum test using the breast cancer dataset**

| | **BA** | **CA** | **CS** | **DA** | **HS** | **ABC** | **FPA** | **MFO** | **PSO** | **BBPSO** | **ThBPSO** |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **ESEM KNN Wilcoxon 10F** | 1.60E-09 | 1.60E-07 | 4.26E-09 | 2.26E-08 | 7.34E-11 | 2.63E-10 | 2.09E-06 | 4.41E-08 | 4.11E-07 | 4.41E-08 | 8.43E-10 |
| **ESEM KNN Wilcoxon HO** | 8.84E-06 | 5.85E-04 | 4.91E-06 | 3.85E-04 | 1.22E-05 | 1.00E-05 | 4.80E-03 | 6.26E-05 | 3.27E-03 | 6.26E-05 | 1.22E-05 |
| **ESEM SVM Wilcoxon 10F** | 4.22E-10 | 2.79E-11 | 4.20E-08 | 4.94E-07 | 5.49E-05 | 3.61E-09 | 3.36E-06 | 1.44E-08 | 6.87E-08 | 1.44E-08 | 1.90E-07 |
| **ESEM SVM Wilcoxon HO** | 7.25E-09 | 2.37E-08 | 1.79E-07 | 1.35E-04 | 2.09E-02 | 1.79E-07 | 1.98E-07 | 3.80E-06 | 3.80E-06 | 3.80E-06 | 1.31E-03 |

| | **GPSO** | **MPSO** | **GMPSO** | **F-BPSO** | **ELPSO** | **MFOPSO** | **AGPSO** | **FS-BPSO** | **DNLPSO** | **BBPSOV** |
|---|---|---|---|---|---|---|---|---|---|---|
| **ESEM KNN Wilcoxon 10F** | 9.97E-11 | 1.61E-04 | 4.92E-10 | 8.33E-09 | 5.41E-09 | 1.12E-07 | 1.01E-10 | 2.29E-06 | 2.90E-11 | 6.91E-11 |
| **ESEM KNN Wilcoxon HO** | 8.08E-06 | 3.18E-02 | 8.08E-06 | 6.84E-06 | 2.22E-05 | 1.89E-03 | 6.36E-06 | 7.64E-06 | 1.23E-11 | 8.08E-06 |
| **ESEM SVM Wilcoxon 10F** | 2.86E-08 | 1.97E-09 | 7.59E-08 | 1.55E-08 | 4.94E-07 | 1.42E-08 | 2.85E-08 | 1.67E-09 | 3.33E-10 | 1.60E-07 |
| **ESEM SVM Wilcoxon HO** | 5.80E-05 | 2.35E-07 | 6.18E-04 | 3.80E-07 | 1.51E-03 | 5.36E-09 | 5.59E-05 | 6.45E-08 | 3.34E-11 | 5.88E-08 |

**Table 6.13: The *p-values* of the Wilcoxon rank sum test using PH2 dataset**

| | ACPSO | BA | CA | CS | DA | HS | ABC | FPA | MFO | PSO | BBPSO |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Average no. of selected features** | 29.25 | 23.64 | 24.66 | 27.54 | 30.21 | 26.74 | 33.50 | 33.81 | 37.99 | 38.65 | 32.09 |
| **Training cost (seconds)** | 2494.46 | 3491.83 | 3016.24 | 2715.68 | 3029.77 | 3130.92 | 3274.45 | 2531.21 | 3397.64 | 3069.28 | 3402.85 |

| | GPSO | MPSO | GMPSO | FB-PSO | ELPSO | MFOPSO | AGPSO | FS-BPSO | DNLPSO | ThBPSO | BBPSOV |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Average no. of selected features** | 28.02 | 24.59 | 28.29 | 26.59 | 24.57 | 24.54 | 33.41 | 30.68 | 29.29 | 31.21 | 35.23 |
| **Training cost (seconds)** | 2507.85 | 2920.20 | 3353.03 | 3033.77 | 3173.42 | 3051.48 | 2517.97 | 2927.54 | 3392.96 | 2968.32 | 2539.79 |

## 6.4 Chapter Summary

In this study, another intelligent system of skin lesion categorisation is presented which consists of ABCD + GLRLM, LBP and HOG feature extraction, ACPSO feature selection, as well as ensemble-based classification. To achieve global optima, the ACPSO model comprises a global search based on increasing and decreasing acceleration coefficients as well as an inclusive local search based on the sub-dimension strategy. In future directions, the assessment of the proposed PSO models can be conducted using other medical image datasets. Moreover, to determine in greater depth how efficient the resulting models is, another possible direction is to enhance deep network structures using the proposed model.

# CHAPTER 7
# PROPOSED METHODOLOGY 5 WITH RCPSO BASED FEATURE SELECTION

We present the modified PSO model, i.e. ACPSO in Chapter 6 where the partial circle, waveform and helix shapes are used for the generation of descending and ascending coefficients. ACPSO has shown a promising advantage in terms of performance and speed in feature selection compared with other conventional and modified PSO methods. In this chapter, the primary focus is to make a further improvement based on the work presented in Chapter 6, by fine-tuning the parameters and restructuring adaptive functions to improve feature selection performance. Specifically, instead of using the partial waveforms, the entire waveforms of the circle, sine and helix functions will be used to generate both coefficients, *C1* and *C2,* to increase search diversity. Instead of using adaptive coefficients as in the previous chapter, RCPSO used random coefficients to gain more search diversity, therefore having better probabilities in achieving the global optima.

## 7.1 The Proposed RCPSO Search Strategies for Feature Selection

We name this new enhanced PSO model as random coefficient PSO (RCPSO). The RCPSO model will start by initiating 50 particles; then the iteration will begin by sorting the swarm based on the population fitness. Three remote leaders are identified using the same process indicated in ACPSO, i.e. after identifying the global best solution, two other leaders with similar fitness scores but remote in positions to the best leader are identified.

The correlation coefficient is utilised for the selection of the second and third leaders. The overall swarm is then split into three sub-swarms randomly. Each sub-swarm is led by one leader; The full waveforms of the circle, sine and helix functions will be generated randomly. 60 values are randomly selected from each waveform. In each sub-swarm, 10 values are randomly selected among the generated 60 values for each waveform and assigned to both coefficients. Therefore, ten offspring solutions are generated by using the 10 randomly selected coefficients in each sub-swarm based search. The best offspring among the 10 new solutions is used to replace the current particle if it is fitter. Three sub-swarms are then combined after a number of iterations. The $gBest$ is updated with the best solution among the swarm. The pseudo-code of RCPSO is shown in Algorithm 6.

**Algorithm 7.1: Pseudo-code of the Proposed RCPSO Model**

| | |
|---|---|
| **1** | Start |
| **2** | Initialize a population randomly (e.g. with 50 Particles); |
| **3** | Evaluate the population to identify the initial best leader**,** $gBest$; |
| **4** | |
| **5** | **While** (! Stop condition) do {// 500 iterations |
| **6** | Find second & third swarm leaders with similar fitness values but low correlation in position to $gBest$; |
| **7** | Divide the population into three sub-swarms, i.e. SS1, SS2 & SS3, with each sub-swarm led by one leader; |
| **8** | **While** (! Stop condition) do {// for three sub-swarms |
| **9** | **For** (each particle $x_i$ in sub-swarm 1) **do** { |
| **10** | Randomly generate 60 values using Eq. 7.1 and randomly select 10 values among the 60 values and assign them to both coefficients respectively. |
| **11** | Generate 10 offspring with Eq. 4.1 & 4.2 using the above randomly assigned coefficients; |
| **12** | Evaluate all offspring and select the best offspring to update $x_i$ if it has better fitness; |
| **13** | **} End For** |
| **14** | Update the sub-warm leader in sub-swarm 1; |
| **15** | **For** (each particle $x_i$ in sub-swarm 2) **do** { |
| **16** | Randomly generate 60 values using Eq. 7.2 and randomly select 10 values among the 60 values and assign them to both coefficients respectively. |
| **17** | Generate 10 offspring with Eq. 4.1 & 4.2 using the randomly selected coefficients; |
| **18** | Evaluate all offspring and select the best offspring to update $x_i$ if it is better; |
| **19** | **} End For** |
| **20** | Update the sub-swarm leader in sub-swarm 2; |
| **21** | **For** (each particle $x_i$ in sub-swarm 3) **do** { |
| **22** | Randomly generate 60 values using Eq. 7.3 and randomly select 10 values among the 60 values and assign them to both coefficients respectively. |
| **23** | Generate 10 offspring with Eq. 4.1 & 4.2 using the randomly selected coefficients; |
| **24** | Evaluate all offspring and select the best offspring to update $x_i$ if it is better; |
| **25** | **} End For** |

| | |
|---|---|
| **26** | Update the sub-swarm leader in sub-swarm 3; |
| **27** | **} End While** |
| **28** | Combine all three sub-swarms; |
| **29** | Sort the population and replace $gBest$ with the new best leader if it is fitter; |
| **30** | **} End While** |
| **31** | Output $gBest$; |
| **32** | **End** |

### 7.1.1 Random Coefficient Generation Using the Full Circle Waveform

In the previous chapter, ACPSO has only utilised the upper right quarter of the circle contour for the adaptive coefficient generation, which might have less diversity. The major difference between ACPSO and this new PSO model is that it utilises the full waveforms including both positive and negative contours to generate the random coefficients and increase search diversity. The circle function defined in Equation (7.1) has been used in sub-swarm 1 to generate the random coefficients.

$$y = \sqrt{r^2 - x^2} \qquad x \in [-2.5, 2.5] \tag{7.1}$$

Equation (7.1) was first used to generate $j$ amount of values in each iteration. These $j$ values are generated randomly, which is sufficient enough to form a full circle shape. The $j$ = 60 has been chosen based on trial and error to represent the full circle waveform; if the $j$ value is less or more than 60, i.e. 50 or 70, the performance of the entire exploration will be slightly affected which may provide less sufficient or over-crowded processes for coefficient generation. In each iteration, the RCPSO model randomly selects 10 values from the above randomly generated 60 values representing the full circle contour, as illustrated in Figure 7-1. RCPSO then assigns the 10 randomly selected values to both coefficients respectively to generate 10 offspring solutions. The best offspring among the 10 new solutions is used to replace the current particle if it has better fitness.

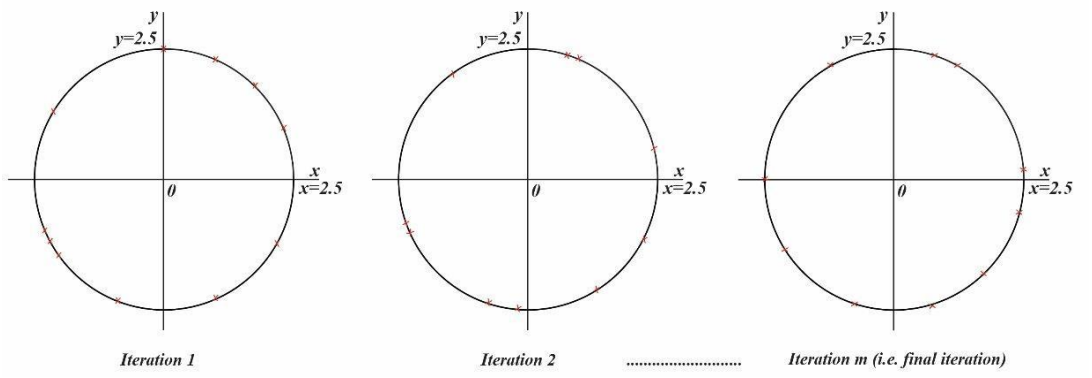



**Figure 7-1: The circle contour for the generation of random coefficients**

### 7.1.2 Random Coefficient Generation Using the Full Sine Waveform

Similar to the above circle-based random coefficient generation, the full sine waveform is also used. A set of 60 positive and negative values was randomly generated by Equation (7.2) for sine waveform in each iteration. Instead of using ascending and descending coefficients, these randomly generated positive and negative values are used to assign both coefficients to guide the search in sub-swarm 2.

$$y = 2.5 \times \sin(x) \qquad x \in [-\pi, \pi] \tag{7.2}$$

These 60 randomly generated values are sufficient enough to form an approximate full sine waveform as shown in Figure 7-2. A total of 10 values are randomly selected among the above generated 60 values in each iteration. They are assigned to both coefficients respectively to generate 10 new offspring solutions. The best solution among the 10 offspring is chosen to replace the current particle if it has better fitness.

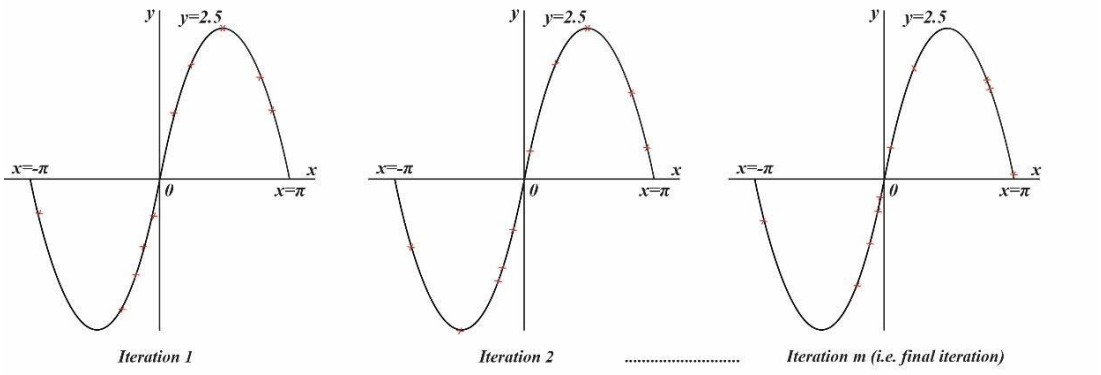



**Figure 7-2: The sine waveform for the generation of random coefficients**

### 7.1.3 Random Coefficient Generation Using the Full Helix Waveform

The above process is also used for the random coefficient generation using the full helix waveform. Specifically, in each iteration, Equations (7.3) -(7.5) are used to generate 60 values randomly to form a full helix contour. Both positive and negative randomly generated values are used to assign both coefficients in sub-swarm 3.

$$x = 0.0065 \times t \times \cos(2 \times t) \qquad x \in [-2.522, 2.532] \tag{7.3}$$

$$y = 0.0065 \times t \times \sin(2 \times t) \qquad y \in [-2.517, 2.517] \tag{7.4}$$

$$z = t \tag{7.5}$$

**where** $t \in [0, 389.5575]$.

A set of 60 positive and negative values is randomly generated using the above equations, which is sufficient enough to form a full helix waveform, as shown in Figure 7-3. Then 10 values among the above generated 60 values are randomly selected and assigned to both coefficients respectively. Then 10 offspring solutions are generated with best one used to replace the current particle if it is fitter.

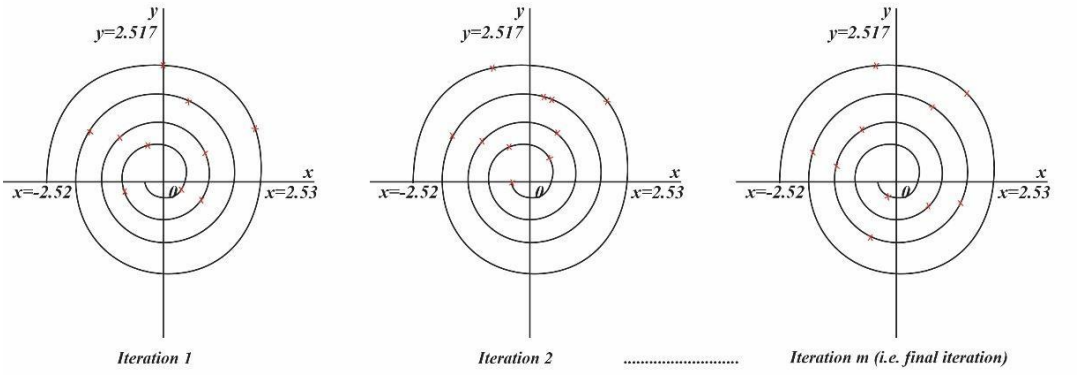



**Figure 7-3: The helix waveform for the generation of random coefficients**

## 7.2 Evaluation

The same combined evaluation dataset (consisting of PH2 and Dermofit) listed in Table 6.1 has also been used in this experiment. A total of 120 images from PH2 and 364 from Dermofit are used for evaluation. An 80:20 aspect ratio for training and testing split is also utilised. Since the proposed RCPSO model did not use any addition function evaluation in each iteration, it has used a default number of function evaluations, i.e. 50 population × 500 iterations × 1 function evaluation = 25,000 function evaluations, as the stopping criterion.

**Table 7.1: Training and testing sets for skin cancer classification**

| Dataset | Benign | Cancer |
|---|---|---|
| **PH2** | 80 | 40 |
| **Dermofit** | 190 | 174 |

### 7.2.1 Parameter Settings

As expansions of PSO, the parameters such as acceleration coefficients, weight and velocity must be altering continuously to acquire an improved performance. The best parameter settings for RCPSO along with other optimisation methods have been identified. Furthermore, experimental and hypothetical research has shown that proposed PSO convergence actions substantially depend on the specific parameters. The RCPSO model utilises an inertia weight of 0.65 to identify the effect of the prior speed on repetitive searches. Random coefficients, C1 and C2, generated by the abovementioned three functions, are employed to expedite convergence. For other search methods, the

parameter settings are employed as stipulated in the previous chapters in sections 4.3.1, 5.2.1 and 6.3.1.

### 7.2.2 Evaluation Using the Combined Dataset

To identify the most suitable feature subsets for each algorithm, a combined dataset made up of both benign and cancerous images will serve as the basis for a 30-run assessment, as shown in Table 7.3 the average GM results attained by the proposed approach integrated with KNN-based and SVM-based ensembles for both 10-fold, and hold-out validations, are higher than those of the other algorithms. RCPSO performs better than other algorithms, as confirmed by the results of the KNN-based ensemble combined with 10-fold validation. Specifically, it outperforms BA, CA, CS, DA, HS, ABC, FPA, MFO, PSO, GPSO, MPSO, GMPSO, FB-PSO, ELPSO, MFOPSO, AGPSO, FS-BPSO, DNLPSO and BBPSOV, respectively. Furthermore, for KNN-based ensemble combined with the hold-out validation, RCPSO outperforms BA, CA, CS, DA, HS, ABC, FPA, MFO, PSO, GPSO, MPSO, GMPSO, FB-PSO, ELPSO, MFOPSO, AGPSO, FS-BPSO, DNLPSO and BBPSOV, respectively. When the SVM-based ensemble classifier as well as 10-fold validation was applied, the proposed algorithm outperforms BA, CA, CS, DA, HS, ABC, FPA, MFO, PSO, GPSO, MPSO, GMPSO, FB-PSO, ELPSO, MFOPSO, AGPSO, FS-BPSO, DNLPSO and BBPSOV, respectively. And lastly, when the SVM-based ensemble classifier combined with the hold-out validation is used, our model outperforms BA, CA, CS, DA, HS, ABC, FPA, MFO, PSO, GPSO, MPSO, GMPSO, FB-PSO, ELPSO, MFOPSO, AGPSO, FS-BPSO, DNLPSO and BBPSOV, respectively. Statistical results illustrated in Table 7.2 also indicate the statistical superiority of the proposed model over other methods.

**Table 7.2: Average performance correlation of each algorithm over 30 runs for the combined dataset**

| | **ESEM KNN** | | **ESEM SVM** | |
|---|---|---|---|---|
| | **10-fold** | **Hold-out** | **10-fold** | **Hold-out** |
| **BA** | 2.17% | 2.31% | 1.41% | 1.45% |
| **CA** | 2.71% | 2.83% | 1.24% | 1.29% |
| **CS** | 2.49% | 2.90% | 0.94% | 1.01% |
| **DA** | 1.75% | 2.27% | 1.03% | 1.13% |
| **HS** | 2.10% | 2.45% | 0.97% | 1.33% |
| **ABC** | 1.98% | 1.92% | 0.76% | 0.89% |
| **FPA** | 2.64% | 2.71% | 1.34% | 1.49% |
| **MFO** | 2.09% | 2.40% | 1.00% | 1.29% |

| | | | | |
|---|---|---|---|---|
| **PSO** | 2.43% | 2.51% | 1.21% | 1.23% |
| **BBPSO** | 2.26% | 2.36% | 1.02% | 1.06% |
| **ThBPSO** | 2.17% | 2.37% | 0.61% | 0.85% |
| **GPSO** | 2.19% | 2.58% | 0.84% | 0.93% |
| **MPSO** | 2.57% | 2.70% | 1.00% | 0.98% |
| **GMPSO** | 2.60% | 2.76% | 0.95% | 1.21% |
| **FB-PSO** | 1.40% | 2.01% | 0.74% | 0.80% |
| **ELPSO** | 1.66% | 2.03% | 0.75% | 0.97% |
| **MFOPSO** | 2.29% | 2.48% | 1.37% | 1.58% |
| **AGPSO** | 2.39% | 2.50% | 1.06% | 1.22% |
| **FS-BPSO** | 1.79% | 2.18% | 0.73% | 1.08% |
| **DNLPSO** | 2.27% | 2.80% | 1.36% | 1.29% |
| **BBPSOV** | 2.63% | 3.03% | 0.82% | 1.04% |

**Table 7.3: Average performance of each algorithm over 30 runs for the combined dataset**

| | **ESEM KNN** | | **ESEM SVM** | |
|---|---|---|---|---|
| | **10-fold** | **Hold-out** | **10-fold** | **Hold-out** |
| **BA** | 0.9362 | 0.9328 | 0.9825 | 0.9809 |
| **CA** | 0.9308 | 0.9276 | 0.9842 | 0.9825 |
| **CS** | 0.9330 | 0.9269 | 0.9872 | 0.9853 |
| **DA** | 0.9404 | 0.9332 | 0.9863 | 0.9841 |
| **HS** | 0.9369 | 0.9314 | 0.9869 | 0.9821 |
| **ABC** | 0.9381 | 0.9367 | 0.9890 | 0.9865 |
| **FPA** | 0.9315 | 0.9288 | 0.9832 | 0.9805 |
| **MFO** | 0.9370 | 0.9319 | 0.9866 | 0.9825 |
| **PSO** | 0.9336 | 0.9308 | 0.9845 | 0.9831 |
| **BBPSO** | 0.9353 | 0.9323 | 0.9864 | 0.9848 |
| **ThBPSO** | 0.9362 | 0.9322 | 0.9905 | 0.9869 |
| **GPSO** | 0.9360 | 0.9301 | 0.9882 | 0.9861 |
| **MPSO** | 0.9322 | 0.9289 | 0.9866 | 0.9856 |
| **GMPSO** | 0.9319 | 0.9283 | 0.9871 | 0.9833 |
| **FB-PSO** | 0.9439 | 0.9358 | 0.9892 | 0.9874 |
| **ELPSO** | 0.9413 | 0.9356 | 0.9891 | 0.9857 |
| **MFOPSO** | 0.9350 | 0.9311 | 0.9829 | 0.9796 |
| **AGPSO** | 0.9340 | 0.9309 | 0.9860 | 0.9832 |
| **FS-BPSO** | 0.9400 | 0.9341 | 0.9893 | 0.9846 |
| **DNLPSO** | 0.9352 | 0.9279 | 0.9830 | 0.9825 |
| **BBPSOV** | 0.9316 | 0.9256 | 0.9884 | 0.9850 |
| **RCPSO** | **0.9579** | **0.9559** | **0.9966** | **0.9954** |

**Table 7.4: The *p-values* of the Wilcoxon rank sum test for the combined dataset**

| | **BA** | **CA** | **CS** | **DA** | **HS** | **ABC** | **FPA** | **MFO** | **PSO** | **BBPSO** |
|---|---|---|---|---|---|---|---|---|---|---|
| **ESEM KNN Wilcoxon 10F** | 7.94E-06 | 3.36E-07 | 5.73E-08 | 1.47E-05 | 3.63E-08 | 3.43E-06 | 2.77E-07 | 1.35E-07 | 6.24E-08 | 1.65E-07 |
| **ESEM KNN Wilcoxon HO** | 7.40E-07 | 3.11E-09 | 2.09E-09 | 6.72E-10 | 5.54E-09 | 1.10E-06 | 4.98E-09 | 3.36E-07 | 2.41E-08 | 4.44E-07 |
| **ESEM SVM Wilcoxon 10F** | 7.16E-06 | 8.03E-06 | 2.18E-04 | 3.00E-07 | 1.52E-04 | 8.68E-04 | 5.01E-06 | 6.98E-07 | 1.53E-05 | 5.44E-05 |
| **ESEM SVM Wilcoxon HO** | 7.67E-08 | 1.95E-07 | 9.04E-06 | 1.49E-09 | 3.32E-07 | 1.10E-05 | 6.40E-08 | 1.59E-08 | 9.33E-06 | 8.32E-06 |

| | **ThBPSO** | **GPSO** | **MPSO** | **GMPSO** | **F-BPSO** | **ELPSO** | **MFOPSO** | **AGPSO** | **FS-BPSO** | **DNLPSO** | **BBPSOV** |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **ESEM KNN Wilcoxon 10F** | 1.06E-07 | 2.00E-07 | 1.46E-07 | 6.76E-09 | 4.97E-04 | 1.86E-05 | 1.35E-07 | 1.46E-07 | 1.02E-06 | 9.02E-08 | 1.48E-08 |
| **ESEM KNN Wilcoxon HO** | 1.55E-08 | 4.19E-08 | 7.67E-09 | 2.29E-09 | 5.05E-08 | 6.98E-08 | 3.36E-07 | 7.67E-09 | 8.85E-07 | 5.37E-10 | 2.62E-10 |
| **ESEM SVM Wilcoxon 10F** | 1.29E-02 | 7.46E-04 | 5.72E-06 | 1.36E-05 | 1.08E-04 | 2.06E-05 | 6.98E-07 | 5.72E-06 | 1.18E-03 | 2.28E-06 | 2.62E-04 |
| **ESEM SVM Wilcoxon HO** | 1.67E-04 | 3.23E-05 | 2.47E-07 | 2.47E-07 | 1.09E-04 | 4.32E-08 | 1.59E-08 | 2.47E-07 | 3.59E-05 | 3.64E-07 | 4.36E-06 |

**Table 7.5: The training computational cost and the average number of features selected using the combined dataset**

| | RCPSO | BA | CA | CS | DA | HS | ABC | FPA | MFO | PSO | BBPSO |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Average no. of selected features** | 61.24 | 74.35 | 67.38 | 75.23 | 78.09 | 71.86 | 70.78 | 70.73 | 65.93 | 73.58 | 73.17 |
| **Training cost (seconds)** | 3724.32 | 3816.57 | 3929.63 | 3927.12 | 3813.44 | 3931.89 | 3906.70 | 3806.12 | 3905.52 | 3878.16 | 3937.50 |

| | GPSO | MPSO | GMPSO | FB-PSO | ELPSO | MFOPSO | AGPSO | FS-BPSO | DNLPSO | ThBPSO | BBPSOV |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Average no. of selected features** | 72.37 | 59.62 | 74.26 | 70.29 | 77.33 | 78.54 | 65.31 | 75.60 | 69.73 | 53.13 | 64.67 |
| **Training cost (seconds)** | 3945.48 | 3998.31 | 3878.48 | 3872.37 | 3970.20 | 3932.88 | 3879.13 | 4011.23 | 3860.92 | 4048.19 | 3981.04 |

### 7.2.3 Evaluation Using the PH2 Dataset

The experiments have also been conducted using the PH2 dataset. A set of 30 runs is also conducted for each method. The detailed evaluation results are shown in Table 7.7, where the best result for each row is highlighted in bold. In combination with both ensemble models under 10-fold and hold-out validations, RCPSO obtains higher average GM performances compared to those of other algorithms. For KNN-based ensemble with 10-fold validation, our results are better than all other methods and outperforms those of BA, CA, CS, DA, HS, ABC, FPA, MFO, PSO, GPSO, MPSO, GMPSO, FB-PSO, ELPSO, MFOPSO, AGPSO, FS-BPSO, DNLPSO and BBPSOV, respectively. Furthermore, when integrated with the KNN-based ensemble and the hold-out validation, it outperforms BA, CA, CS, DA, HS, ABC, FPA, MFO, PSO, GPSO, MPSO, GMPSO, FB-PSO, ELPSO, MFOPSO, AGPSO, FS-BPSO, DNLPSO and BBPSOV, respectively. When the SVM-based ensemble classifier combined with 10-fold validation is employed, the proposed algorithm outperforms BA, CA, CS, DA, HS, ABC, FPA, MFO, PSO, GPSO, MPSO, GMPSO, FB-PSO, ELPSO, MFOPSO, AGPSO, FS-BPSO, DNLPSO and BBPSOV, respectively. When the SVM-based ensemble and the hold-out validation are used, our model outperforms BA, CA, CS, DA, HS, ABC, FPA, MFO, PSO, GPSO, MPSO, GMPSO, FB-PSO, ELPSO, MFOPSO, AGPSO, FS-BPSO, DNLPSO and BBPSOV, respectively. As indicated by the statistical results in Table 7.4, RCPSO shows significant improvements over nearly all other methods for skin lesion classification. Again, FS-BPSO shows the best performances among the baseline methods.

**Table 7.6: Average performance correlation of each algorithm over 30 runs for the PH2 dataset**

| | ESEM KNN | | ESEM SVM | |
|---|---|---|---|---|
| | **10-fold** | **Hold-out** | **10-fold** | **Hold-out** |
| **BA** | 3.35% | 1.98% | 3.14% | 3.99% |
| **CA** | 3.82% | 1.65% | 1.61% | 3.25% |
| **CS** | 3.48% | 2.04% | 1.17% | 2.34% |
| **DA** | 3.79% | 1.59% | 1.86% | 3.45% |
| **HS** | 2.39% | 0.86% | 1.66% | 3.01% |
| **ABC** | 3.57% | 1.92% | 1.51% | 2.68% |
| **FPA** | 4.09% | 2.43% | 2.76% | 3.95% |
| **MFO** | 2.92% | 1.82% | 1.22% | 2.22% |
| **PSO** | 3.45% | 2.11% | 1.61% | 2.82% |
| **BBPSO** | 3.77% | 2.44% | 1.68% | 3.07% |
| **ThBPSO** | 2.55% | 1.14% | 1.87% | 2.49% |
| **GPSO** | 2.86% | 1.36% | 1.45% | 3.08% |
| **MPSO** | 3.28% | 1.76% | 2.50% | 2.76% |
| **GMPSO** | 3.65% | 2.05% | 2.02% | 3.42% |
| **FB-PSO** | 3.52% | 1.25% | 1.63% | 3.19% |
| **ELPSO** | 3.57% | 1.73% | 1.64% | 3.22% |
| **MFOPSO** | 3.75% | 2.51% | 1.98% | 2.26% |
| **AGPSO** | 3.87% | 2.39% | 1.33% | 2.30% |
| **FS-BPSO** | 2.12% | 0.55% | 1.33% | 2.73% |
| **DNLPSO** | 2.71% | 1.77% | 1.75% | 2.86% |
| **BBPSOV** | 3.70% | 1.93% | 1.37% | 2.54% |

**Table 7.7: Average performance of each algorithm over 30 runs for the PH2 dataset**

| | ESEM KNN | | ESEM SVM | |
|---|---|---|---|---|
| | **10-fold** | **Hold-out** | **10-fold** | **Hold-out** |
| **BA** | 0.9407 | 0.9534 | 0.9465 | 0.9355 |
| **CA** | 0.9360 | 0.9567 | 0.9618 | 0.9429 |
| **CS** | 0.9394 | 0.9528 | 0.9662 | 0.9520 |
| **DA** | 0.9363 | 0.9573 | 0.9593 | 0.9409 |
| **HS** | 0.9503 | 0.9646 | 0.9613 | 0.9453 |
| **ABC** | 0.9385 | 0.9540 | 0.9628 | 0.9486 |
| **FPA** | 0.9333 | 0.9489 | 0.9503 | 0.9359 |
| **MFO** | 0.9450 | 0.9550 | 0.9657 | 0.9532 |
| **PSO** | 0.9397 | 0.9521 | 0.9618 | 0.9472 |
| **BBPSO** | 0.9365 | 0.9488 | 0.9611 | 0.9447 |
| **ThBPSO** | 0.9487 | 0.9618 | 0.9592 | 0.9505 |
| **GPSO** | 0.9456 | 0.9596 | 0.9634 | 0.9446 |
| **MPSO** | 0.9414 | 0.9556 | 0.9529 | 0.9478 |
| **GMPSO** | 0.9377 | 0.9527 | 0.9577 | 0.9412 |
| **FB-PSO** | 0.9390 | 0.9607 | 0.9616 | 0.9435 |
| **ELPSO** | 0.9385 | 0.9559 | 0.9615 | 0.9432 |
| **MFOPSO** | 0.9367 | 0.9481 | 0.9581 | 0.9528 |
| **AGPSO** | 0.9355 | 0.9493 | 0.9646 | 0.9524 |
| **FS-BPSO** | 0.9530 | 0.9677 | 0.9646 | 0.9481 |
| **DNLPSO** | 0.9471 | 0.9555 | 0.9604 | 0.9468 |
| **BBPSOV** | 0.9372 | 0.9539 | 0.9642 | 0.9500 |
| **RCPSO** | **0.9742** | **0.9732** | **0.9779** | **0.9754** |

**Table 7.8: The *p*-values of the Wilcoxon rank sum test for the PH2 Dataset**

| | BA | CA | CS | DA | HS | ABC | FPA | MFO | PSO | BBPSO |
|---|---|---|---|---|---|---|---|---|---|---|
| **ESEM KNN Wilcoxon 10F** | 4.15E-08 | 1.23E-08 | 2.83E-06 | 4.39E-08 | 4.24E-04 | 2.24E-07 | 1.28E-06 | 3.66E-08 | 1.49E-06 | 2.19E-05 |
| **ESEM KNN Wilcoxon HO** | 2.69E-09 | 5.40E-05 | 1.20E-06 | 1.36E-06 | 1.50E-03 | 1.62E-05 | 1.28E-08 | 6.69E-08 | 1.13E-06 | 1.42E-08 |
| **ESEM SVM Wilcoxon 10F** | 5.93E-06 | 1.14E-03 | 1.16E-02 | 1.10E-04 | 6.95E-04 | 6.59E-04 | 5.18E-05 | 4.81E-05 | 4.31E-04 | 5.74E-04 |
| **ESEM SVM Wilcoxon HO** | 1.27E-07 | 8.46E-07 | 7.49E-04 | 8.48E-08 | 2.76E-06 | 6.58E-05 | 1.95E-08 | 1.31E-04 | 5.47E-05 | 1.32E-05 |

| | ThBPSO | GPSO | MPSO | GMPSO | F-BPSO | ELPSO | MFOPSO | AGPSO | FS-BPSO | DNLPSO | BBPSOV |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **ESEM KNN Wilcoxon 10F** | 1.51E-05 | 5.60E-06 | 8.32E-07 | 1.69E-09 | 5.31E-06 | 9.62E-07 | 3.66E-08 | 8.32E-07 | 3.49E-02 | 1.87E-06 | 3.86E-08 |
| **ESEM KNN Wilcoxon HO** | 1.29E-03 | 4.74E-04 | 4.24E-10 | 6.66E-08 | 2.77E-03 | 1.44E-08 | 6.69E-08 | 4.24E-10 | **3.93E-01** | 4.85E-06 | 1.49E-08 |
| **ESEM SVM Wilcoxon 10F** | 9.33E-05 | 1.09E-03 | 1.09E-02 | 1.54E-04 | 7.15E-03 | 1.79E-03 | 4.81E-05 | 1.09E-02 | **5.69E-02** | 6.01E-05 | 1.15E-02 |
| **ESEM SVM Wilcoxon HO** | 4.70E-04 | 9.04E-07 | 3.01E-04 | 1.84E-06 | 2.91E-07 | 6.53E-07 | 1.31E-04 | 3.01E-04 | 1.39E-05 | 5.61E-06 | 1.80E-04 |

**Table 7.9: The training computational cost and the average number of features selected using PH2 dataset**

| | RCPSO | BA | CA | CS | DA | HS | ABC | FPA | MFO | PSO | BBPSO |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Average no. of selected features** | 64.32 | 81.57 | 75.08 | 78.41 | 83.13 | 73.89 | 78.12 | 75.45 | 69.65 | 83.61 | 82.01 |
| **Training cost (seconds)** | 3697.25 | 3823.17 | 3933.10 | 3937.52 | 3817.28 | 3933.13 | 3914.06 | 3811.73 | 3915.43 | 3885.54 | 3942.80 |

| | GPSO | MPSO | GMPSO | FB-PSO | ELPSO | MFOPSO | AGPSO | FS-BPSO | DNLPSO | ThBPSO | BBPSOV |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Average no. of selected features** | 73.46 | 66.76 | 80.18 | 70.68 | 85.61 | 84.20 | 71.35 | 78.71 | 71.01 | 54.38 | 74.79 |
| **Training cost (seconds)** | 3950.32 | 4005.86 | 3881.97 | 3876.67 | 3976.87 | 3935.33 | 3881.85 | 4020.29 | 3863.69 | 4054.51 | 3981.54 |

**Table 7.10: Performance comparison with related research for the PH2 database**

| Studies | Methodology | Classes | Evaluation Strategy | Recognition rate (%) |
|---|---|---|---|---|
| **(Abuzaghleh et al., 2014b)** | 2-D Fast Fourier Transform features set + 2-D Discrete Cosine Transform features set + k-Nearest Neighbour | 3 | 25% for testing | 63.33 |
| **(Abuzaghleh et al., 2015b)** | 2-D Fast Fourier Transform + 2-D Discrete Cosine Transform + Complexity + Colour + Pigment Network Feature Set + No Classification Mentioned | 3 | 5-fold | 96.50 |
| **(Barata et al., 2012)** | Lesion or pigment ratio features + Boosting algorithm | 3 | Ten-Fold | 86.20 |
| **(Barata et al., 2013)** | Colour Scale Invariant Feature Transform + BoF + No Classification Mentioned | 3 | Ten-Fold | 87.00 |
| **(Barata et al., 2015)** | Colour Constancy + BoF Framework + k-mean + Support vector machine | 3 | Ten-Fold | 84.30 |
| **(Soumya et al., 2016)** | Colour Correlogram + Segmentation based Fractal Texture Analysis + Bayes Classifier | 3 | Ten-Fold | 91.50 |
| **(Pennisi et al., 2016)** | Shape Features using artefact removal, Skin detection, lesion segmentation & binary mask (ASLM) + AdaBoost | 3 | N/A | 93.60 |
| **(Vasconcelos et al., 2015)** | Colour Features + Support vector machine | 3 | Ten-Fold | 81.38 |
| **(Waheed et al., 2017)** | Colour (HSV)+ GLCM Texture + SVM | 3 | 3-folds | 96.00 |
| **(Marques et al., 2012)** | Texture + Colour (RGB, HSV, Lab) + global feature vector (histogram) | 3 | N/A | 79.10 |

| | | | | |
|---|---|---|---|---|
| **(Bi et al., 2016)** | Multi-scale lesion-biased representation + Joint Reverse Classification + SVM with linear kernel | 3 | N/A | 92.00 |
| **(Riaz et al., 2014)** | LBP + Colour (HSV) + Principal component analysis + SVM | 3 | 10-folds | 89.00 |
| **(Alfed et al., 2016)** | Colour Histograms + Colour moments + HOG + Codebook generation + SVM, KNN & AdaBoost | 3 | 5-folds | 88.00 |
| **(Eltayef et al., 2017)** | Properties of pigment network + ANN | 3 | N/A | 90.00 |
| **(Alfed and Khelifi, 2017)** | SIFT, HOG, Histogram of Oriented Lines + Colour Vector Angles + Zernike Moments + Codebook generation + | 3 | 5-folds | 98.79 |
| **(Bi et al., 2017)** | Multistage Fully Convolutional Networks | 3 | N/A | 94.24 |
| **(Jamil et al., 2018)** | Colour Mean+ Colour moments + Colour auto correlogram + Grey intensity-based features + Gabor wavelet + Shape + Local fisher discriminant analysis +SVM | 3 | N/A | 98.29 |
| **GA (Chapter 3)** | **Shape, colour & GLCM texture features + GA + SVM** | **3** | **Ten-Fold** | **92.11** |
| **PSO (Chapter 3)** | **Shape, colour & GLRLM texture features + PSO + SVM** | **3** | **Ten-Fold** | **90.93** |
| **HLPSO (Chapter 4)** | **Shape, colour & GLRLM texture features + HLPSO + SVM** | **3** | **Ten-Fold and Hold Out** | **94.32 96.14** |
| **PSOVA (Chapter 5)** | **Shape, colour & GLRLM texture features + PSOVA + SVM** | **3** | **Ten-Fold and Hold Out** | **95.23 96.45** |
| **ACPSO (Chapter 6)** | **Shape, colour & GLRLM texture features + LBP + HOG + ACPSO + ESEM-SVM** | **3** | **Ten-Fold and Hold Out** | **97.67 97.31** |
| **RCPSO (Chapter 7)** | **Shape, colour & GLRLM texture features + LBP + HOG + RCPSO + ESEM-SVM** | **3** | **Ten-Fold and Hold Out** | **97.79 97.54** |

This research is comparatively analysed with other relevant studies on the detection of skin cancer based on the PH2 dataset in Table 7.10. An estimation of performance can be obtained as every study employed distinctive training and testing sets and each of them was assessed based on different evaluation strategies. The modified PSO models proposed in the present study are among the top performers and provide the most effective alternative options for the classification of skin lesions.

### 7.2.4 Evaluation Using the UCI Epileptic Seizure Dataset

For the assessment of how the proposed RCPSO model performs for feature optimization, an extra, more extensive scale dataset was employed, namely, the UCI epileptic seizure dataset. This dataset includes a total of 11,500 samples, 178 features and five categories. Since this is a large dataset, only a randomly selected subset was chosen for evaluation, which consists of 6,000 data. The experimental set-up for the assessment involved an 80:20 training-to-testing aspect ratio. For each method, the following experimental setting is applied, i.e. 30 runs, 178 dimensions, 20 populations and 200 repetitions, and 20

population × 200 maximum number of repetitions as the maximum number of function assessments. Tables 5.7 and 5.8 provide the average GM score across 30 runs and the outcomes of the Wilcoxon rank sum test respectively. Based on these results, the performance of the proposed RCPSO model are statistically better than those of other alternative approaches for nearly all test cases. The exception is for FS-BPSO, which has similar mean performance distribution to those of RCPSO, when integrated with the SVM-based ensemble and the 10-fold validation. In summary, the proposed RCPSO model has greater efficiency in generating solutions for the selection of discriminative features regarding the epileptic seizure dataset.

**Table 7.11: Average performance of each algorithm over 30 runs for the epileptic seizure dataset**

| | **KNN** | | **SVM** | |
|---|---|---|---|---|
| | **10-fold** | **Hold-out** | **10-fold** | **Hold-out** |
| **BA** | 0.8014 | 0.8074 | 0.9394 | 0.9408 |
| **CA** | 0.8077 | 0.8131 | 0.9385 | 0.9398 |
| **CS** | 0.8033 | 0.8092 | 0.9401 | 0.9413 |
| **DA** | 0.7996 | 0.8052 | 0.9370 | 0.9383 |
| **HS** | 0.8026 | 0.8085 | 0.9374 | 0.9388 |
| **ABC** | 0.8013 | 0.8069 | 0.9376 | 0.9388 |
| **FPA** | 0.8030 | 0.8085 | 0.9376 | 0.9389 |
| **MFO** | 0.8041 | 0.8095 | 0.9364 | 0.9378 |
| **PSO** | 0.8061 | 0.8114 | 0.9380 | 0.9395 |
| **BBPSO** | 0.8041 | 0.8095 | 0.9364 | 0.9378 |
| **ThBPSO** | 0.8024 | 0.8082 | 0.9382 | 0.9395 |
| **GPSO** | 0.8030 | 0.8085 | 0.9401 | 0.9414 |
| **MPSO** | 0.8003 | 0.8058 | 0.9353 | 0.9367 |
| **GMPSO** | 0.7997 | 0.8054 | 0.9357 | 0.9370 |
| **F-BPSO** | 0.8090 | 0.8130 | 0.9410 | 0.9425 |
| **ELPSO** | 0.7961 | 0.8025 | 0.9352 | 0.9365 |
| **MFOPSO** | 0.8046 | 0.8104 | 0.9375 | 0.9388 |
| **AGPSO** | 0.7983 | 0.8040 | 0.9375 | 0.9389 |
| **FS-BPSO** | 0.8105 | 0.8135 | 0.9425 | 0.9433 |
| **DNLPSO** | 0.8033 | 0.8089 | 0.9375 | 0.9388 |
| **BBPSOV** | 0.8015 | 0.8068 | 0.9399 | 0.9411 |
| **RCPSO** | **0.8199** | **0.8230** | **0.9478** | **0.9484** |

**Table 7.12: The *p-values* of the Wilcoxon rank sum test using the epileptic seizure dataset**

| | **BA** | **CA** | **CS** | **DA** | **HS** | **ABC** | **FPA** | **MFO** | **PSO** | **BBPSO** | **ThBPSO** |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **KNN Wilcoxon 10F** | 2.38E-07 | 3.56E-04 | 2.68E-06 | 5.46E-09 | 1.73E-07 | 6.05E-07 | 7.04E-07 | 9.79E-05 | 2.77E-05 | 9.79E-05 | 3.57E-06 |
| **KNN Wilcoxon HO** | 4.16E-07 | 7.48E-04 | 5.53E-06 | 1.39E-08 | 3.33E-07 | 6.88E-07 | 9.05E-07 | 3.33E-04 | 5.38E-05 | 3.33E-04 | 1.95E-05 |
| **SVM Wilcoxon 10F** | 3.01E-04 | 3.08E-08 | 2.13E-04 | 6.01E-08 | 1.55E-09 | 1.41E-09 | 2.03E-07 | 4.44E-07 | 6.74E-06 | 4.44E-07 | 5.09E-08 |
| **SVM Wilcoxon HO** | 8.69E-04 | 2.04E-07 | 5.88E-04 | 7.76E-08 | 8.68E-10 | 6.39E-09 | 3.10E-07 | 2.65E-06 | 4.66E-05 | 2.65E-06 | 4.90E-07 |

| | GPSO | MPSO | GMPSO | F-BPSO | ELPSO | MFOPSO | AGPSO | FS-BPSO | DNLPSO | BBPSOV |
|---|---|---|---|---|---|---|---|---|---|---|
| **KNN Wilcoxon 10F** | 4.31E-08 | 2.03E-07 | 9.06E-08 | 1.73E-05 | 2.15E-10 | 1.61E-06 | 3.35E-08 | 2.77E-04 | 5.61E-05 | 6.74E-06 |
| **KNN Wilcoxon HO** | 6.85E-07 | 1.29E-06 | 2.20E-06 | 3.01E-04 | 9.30E-10 | 3.75E-06 | 1.85E-07 | 5.38E-05 | 8.17E-05 | 1.01E-05 |
| **SVM Wilcoxon 10F** | 1.36E-07 | 1.10E-08 | 1.07E-09 | 2.13E-02 | 3.02E-11 | 1.61E-06 | 2.03E-07 | **3.98E-01** | 1.70E-08 | 2.57E-07 |
| **SVM Wilcoxon HO** | 3.07E-07 | 5.01E-08 | 3.05E-09 | 4.10E-02 | 4.66E-11 | 3.84E-06 | 2.64E-07 | 2.13E-02 | 2.71E-08 | 6.76E-07 |

**Table 7.13: The *p-values* of the Wilcoxon rank sum test using PH2 dataset**

| | RCPSO | BA | CA | CS | DA | HS | ABC | FPA | MFO | PSO | BBPSO |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Average no. of selected features** | 87.36 | 84.56 | 86.44 | 85.88 | 89.82 | 92.86 | 95.94 | 97.53 | 95.15 | 91.58 | 98.66 |
| **Training cost (seconds)** | 3598.73 | 4332.90 | 4219.15 | 3706.39 | 4040.26 | 4315.96 | 4049.33 | 3719.27 | 4143.33 | 4032.66 | 3885.63 |

| | GPSO | MPSO | GMPSO | FB-PSO | ELPSO | MFOPSO | AGPSO | FS-BPSO | DNLPSO | ThBPSO | BBPSOV |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Average no. of selected features** | 98.38 | 91.95 | 97.55 | 97.30 | 98.04 | 99.87 | 103.20 | 98.79 | 101.42 | 106.86 | 107.33 |
| **Training cost (seconds)** | 4533.92 | 4043.87 | 4518.30 | 4241.27 | 4506.93 | 4560.74 | 4287.72 | 3762.33 | 4317.28 | 3777.65 | 4386.75 |

## 7.3 Chapter Summary

A RCPSO-based feature selection model is proposed in this study. Instead of using adaptive coefficients as in ACPSO, random acceleration coefficients are generated using the full waveforms of circle, sine and helix functions in RCPSO which enhances search diversification. In future studies, the newly proposed model can be assessed using additional UCI datasets. The optimal parameter selection can also be exploited using the proposed model.

# CHAPTER 8
# CONCLUSION AND FUTURE WORK

## 8.1 Summary of Research Contributions

The skin cancer detection systems proposed in the present study are underpinned by PSO-based feature selection based on novel features of colour and texture, such as the ABCD, LBP, HOG and GLRLM features with effective classification of skin cancer lesions.

**Contribution 1**

The first contribution of this research is to use PSO and GA for feature selection to enhance skin lesion classification performance. When evaluated with 100 skin lesion images, the system has reached with promising classification accuracies of 92.11% for GA and 90.93% for PSO based feature selection respectively.

**Contribution 2**

The second contribution is the proposal of a modified PSO, known as hybrid learning PSO (HLPSO), for feature selection. In order to overcome premature convergence of the original PSO model, HLPSO integrates PSO with multiple strategies, i.e. sub-swarm division, random explorations using Gaussian, Cauchy and Levy distributions, crossover and mutation operators and scattering strategies.

Instead of using the whole swarm for exploration, HLPSO has divided the original population into two sub-swarms. The search process in one sub-swarm is guided by long and short exploration using probability distributions, while the other sub-swarm is led by

the original PSO model. Moreover, the 10 best particles of the overall swarm are selected for further crossover and mutation to create 20 offspring solutions. A scattering mechanism is also used to re-locate the worst particles in the swarm to diversify the population. Overall, the result performance has been significantly improved while compared with other conventional feature selection methods including those used in contribution 1. The proposed method achieves GM performances of 93.88% for KNN combined with 10-fold validation, 95.62% for KNN integrated with hold-out validation, 93.54% for SVM combined with 10-fold validation and 96.00% for SVM incorporated with hold-out validation, respectively.

**Contribution 3**

The third contribution of this research is to put forth another PSO variant model, i.e. the PSOVA algorithm, for feature selection and skin cancer classification.

To prevent the premature convergence of the conventional PSO algorithm, PSOVA includes more diverse velocity updating strategies than those of HLPSO for feature optimization. Specifically, PSOVA combines the following essential steps to diversify the search, i.e. sub-swarm division, attraction and evading actions guided by multiple swarm leasers and worst signals, swarm leader enhancement and diverse matrix representations. In comparison with the classical methods employed for the evaluation of HLPSO and state-of-the-art PSO variants, the PSOVA model is more effective in identifying the most significant features for detecting both benign and cancerous lesions. With an emphasis on the selection of features based on attraction and evading actions, the key mechanisms of PSOVA are discussed below.

To decrease the chances of trapping in local optima, two remote leaders are identified initially which have similar fitness scores but low correlation in positions. In other words, the selected two swarm leaders must assure to have an appropriate distance between each other, and so the search processes could reach different regions. The two remote leaders are then used to guide the sub-swarm based searches.

To diversify the search processes, PSOVA enables the particles to follow local and global best solutions and worst signals in each dimension and randomly selected partial dimensions. The best solution identified in each sub-swarm is also further improved, and the exploitation is augmented by employing the three random walks of Gaussian, Cauchy and Levy distributions.

The search is further diversified using a dynamic matrix representation of the swarm in each generation. In contrast to other search strategies and methods, the proposed PSOVA algorithm is highly efficient not only in choosing ideal features for categorisation of melanomas, but also in addressing unimodal and multimodal benchmark functions. Additionally, as compared with the approaches used in previous studies on skin cancer classification, the algorithm has achieved the highest performances when it comes to detecting skin cancer. In comparison with the results obtained in contribution 2, the experiments using PSOVA for feature selection have increased the average GM performances over 30 runs by 0.28% (94.16%) and 0.43% (96.04%) for KNN with 10-fold and hold-out validations, and 1.69% (95.23%) and 0.45% (96.45%) for SVM with 10-fold and hold-out validations, respectively.

**Contribution 4**

The fourth major contribution of this research is the extraction of diverse lesion features using multiple feature descriptors and the proposal of an adaptive coefficient PSO model, known as ACPSO, for feature selection. Two ensemble classification models are also developed by integrating several base classifiers dedicated to each feature type to enhance performance.

First of all, the clinical diagnosis prompts the initial extraction of a series of texture features (e.g. GLRLM) and the ABCD features, including shape and colour features. To better represent the lesions, other more refined textural features are also extracted using LBP and HOG descriptors, respectively.

A new improved PSO algorithm (ACPSO) is also proposed for feature selection. In comparison with HLPSO and PSOVA, ACPSO employs adaptive acceleration

coefficients to improve the skin lesion detection performance. Besides integrating features for a global search using adaptive coefficients, the ACPSO model also divides features into sub-dimensions to enhance feature optimization for each lesion region (e.g. top, middle and bottom regions of the lesion). The search process of ACPSO begins with the identification of three remote swarm leaders with similar fitness scores but low correlation in positions to guide three sub-swarm based searches, with the search process of each sub-swarm led by one swarm leader.

The ascending and descending acceleration coefficients are proposed in ACPSO. These adaptive coefficients are generated dynamically based on partial circle, sine and helix waveforms to guide sub-swarm based searches. These coefficients adaptation processes ensure the search will focus on global exploration from the start of the iterations and shift to local exploitation in later stage of the iterations to accelerate convergence.

Furthermore, a sub-dimension-based search is also performed to obtain more refined important features to improve lesion classification. Finally, a few new positions are re-initialised randomly to replace of the weak solutions in the population. ACPSO outperforms the previously employed classical search methods and advanced PSO variants for diverse feature selection and unimodal and multimodal optimization problems. It has significantly improved the performances and achieved average GM performances of 97.26% and 97.38% for KNN-based ensemble with 10-fold and hold-out validations, and 97.67% and 97.31% for SVM-based ensemble with 10-fold and hold-out validations, respectively.

**Contribution 5**

The fifth major contribution of this study is to further improve the above proposed ACPSO model by using dynamically generated random coefficients. This improved model is named as random coefficient PSO (RCPSO). Specifically, RCPSO employs dynamic random acceleration coefficients generated based on the full circle, sine and helix waveforms instead of ascending and descending coefficients produced by partial waveforms as in ACPSO.

The RCPSO model uses positive and negative random acceleration coefficients supported by utilising full waveforms resulting from non-linear circle, sine and helix functions to explore the search space. Since the random coefficients generated in each iteration are more diverse including both positive and negative values, RCPSO explores a wider search space, therefore having more chances of finding global optima.

In addition to making it possible to explore the search space more widely by enhancing diversification, random coefficients also fine-tune the areas of local and global best solutions. Hence, the RCPSO model is better equipped to find the global optima because it replicates the hovering flight behaviour of hummingbirds around attraction (e.g. flower or food source) and can investigate a broader search space. Lastly, the RCPSO model makes further performance improvement and achieves average GM performances of 97.42% and 97.32% for KNN-based ensemble model with 10-fold and hold-out validations, and 97.79% and 97.54% for SVM-based ensemble model with 10-fold and hold-out validations, respectively.

## 8.2 Future Directions

In future work, more skin lesion features can be integrated for the experiment. A codebook with effective characterisation of skin cancer lesions based on novel features of colour and texture, such as the third order Zernike moments, the histogram of oriented lines (HOL), and colour vector angles, can be used to extend the classification performance. In addition, several possible directions of this study can be developed in the future. For example, the proposed models of skin lesion detection could be expanded to address the recognition of different cancer stages and occurrence of more than two or three common types of skin cancers.

Despite the use of various public databases, a larger dataset with numerous types and labelled in a more abundant manner with additional dimensions (e.g. additional texture-based features) would facilitate not only extension of the suggested skin cancer recognition schemes, but also adding a more in-depth investigation of how the various lesion features are correlated. In addition, a better skin cancer classification scheme could

potentially incorporate diverse ensemble base models trained on colour and texture features as well as characteristics extracted using other feature extraction methods.

Furthermore, additional improvement of feature optimisation created by different evolutionary algorithms can be considered such as the firefly algorithm and the cuckoo search algorithm. Moreover, to enhance the existing system for addressing challenging real-life optimisation issues, embed the proposed PSO algorithms with multiple objectives could be useful. Additionally, to create brain-computer interfaces and produce a more satisfying user experience, the proposed system could be implemented and operated on additional platforms, including smartphones, and humanoid robots.

Previous studies in the literature (Duro et al., 2012, Tharwat et al., 2017, Yang and Xu, 2017) have indicated that single classifiers (e.g. KNN, SVM) can be made to perform better by using evolutionary optimisation algorithms through determination of finest parameter configurations. In addition, ensemble classifiers can also achieve better classification performance with ensemble construction and optimisation strategies based on evolutionary algorithms (Zhang et al., 2018b). Thus, future research will investigate parameter and structure tuning for both single and ensemble models based on metaheuristic search algorithms to make detection of skin cancer more precise. Meanwhile, deep learning methods constitute an additional developing research field that holds considerable potential for skin lesion detection. A wide range of deep learning configurations have been suggested for various machine vision issues, performing remarkably well in vision-based applications in real life. Thus, given their efficiency, deep learning methods are worth considering, particularly regarding automatic skin cancer detection in real-life environments.

To accomplish with fully automated medical scan systems and thus improve and simplify user experience, it is possible to deploy the proposed models to real-world applications, such as integrating them with smart automation systems or built-in medical devices, for clinical diagnosis. The proposed models can also be used to develop a mobile-based diagnosis application to allow the deployment in remote areas which lack of medical assistance, thus being able to conduct instant and early diagnosis. In future work, more

evaluations will also be conducted using other larger datasets with larger feature dimensions to further evaluate the efficiency of the proposed models. What is more, bioinformatics and other application domains could derive advantages from the suggested methods of feature extraction and selection, which could be expanded without difficulty. For instance, it is possible to tune and implement the proposed algorithms into wider medical domains, to deal with, such as CT or MRI brain tumour, retinal disease and blood cancer classification and detection.

# CHAPTER 9
# REFERENCE


ABBOTT, L. M. & SMITH, S. D. 2018. Smartphone apps for skin cancer diagnosis: Implications for patients and practitioners. *Australasian Journal of Dermatology*.

ABUZAGHLEH, O., BARKANA, B. D. & FAEZIPOUR, M. 2014a. Automated skin lesion analysis based on color and shape geometry feature set for melanoma early detection and prevention. *Systems, Applications and Technology Conference (LISAT), 2014 IEEE Long Island.* IEEE.

ABUZAGHLEH, O., BARKANA, B. D. & FAEZIPOUR, M. SKINcure: A real time image analysis system to aid in the malignant melanoma prevention and early detection. Image Analysis and Interpretation (SSIAI), 2014 IEEE Southwest Symposium on, 2014b. IEEE, 85-88.

ABUZAGHLEH, O., BARKANA, B. D. & FAEZIPOUR, M. 2015a. Noninvasive real-time automated skin lesion analysis system for melanoma early detection and prevention. *IEEE journal of translational engineering in health and medicine,* 3**,** 1-12.

ABUZAGHLEH, O., FAEZIPOUR, M. & BARKANA, B. D. A comparison of feature sets for an automated skin lesion analysis system for melanoma early detection and prevention. Systems, Applications and Technology Conference (LISAT), 2015 IEEE Long Island, 2015b. IEEE, 1-6.

ADJED, F., FAYE, I. & ABABSA, F. Segmentation of skin cancer images using an extension of chan and vese model. Information Technology and Electrical Engineering (ICITEE), 2015 7th International Conference on, 2015. IEEE, 442-447.

AIN, Q. U., XUE, B., AL-SAHAF, H. & ZHANG, M. Genetic programming for skin cancer detection in dermoscopic images. Evolutionary Computation (CEC), 2017 IEEE Congress on, 2017. IEEE, 2420-2427.

AITKEN, J. F., PFITZNER, J., BATTISTUTTA, D., O'ROURKE, P. K., GREEN, A. C. & MARTIN, N. G. 1996. Reliability of computer image analysis of pigmented skin lesions of Australian adolescents. *Cancer,* 78**,** 252-257.

AITKEN, J. F., YOULDEN, D. R., BAADE, P. D., SOYER, H. P., GREEN, A. C. & SMITHERS, B. M. 2018. Generational shift in melanoma incidence and mortality in Queensland, Australia, 1995–2014. *International journal of cancer,* 142**,** 1528-1535.

ALCÓN, J. F., CIUHU, C., TEN KATE, W., HEINRICH, A., UZUNBAJAKAVA, N., KREKELS, G., SIEM, D. & DE HAAN, G. 2009. Automatic imaging system with decision support for inspection of pigmented skin lesions and melanoma diagnosis. *IEEE journal of selected topics in signal processing,* 3**,** 14-25.

ALFED, N. & KHELIFI, F. 2017. Bagged textural and color features for melanoma skin cancer detection in dermoscopic and standard images. *Expert Systems with Applications,* 90**,** 101-110.

ALFED, N., KHELIFI, F. & BOURIDANE, A. Improving a bag of words approach for skin cancer detection in dermoscopic images. Control, Decision and Information Technologies (CoDIT), 2016 International Conference on, 2016. IEEE, 024-027.

ALFED, N., KHELIFI, F., BOURIDANE, A. & SEKER, H. Pigment network-based skin cancer detection. Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE, 2015. IEEE, 7214-7217.

ANDRUSHIA, A. D. & PATRICIA, A. T. J. N. I. O. T. F. I. P. A. 2018. Artificial Bee Colony Based Feature Selection for Automatic Skin Disease. 150**,** 215.

AYOUB, A., HAJDU, A. & NAGY, A. 2012. Automatic detection of pigmented network in melanoma dermoscopic images. *The International Journal of Computer Science and Communication Security (IJCSCS),* 2**,** 58-63.

BACK, T. 1996. *Evolutionary algorithms in theory and practice: evolution strategies, evolutionary programming, genetic algorithms*, Oxford university press.

BALLERINI, L., FISHER, R. B., ALDRIDGE, B. & REES, J. Non-melanoma skin lesion classification using colour image data in a hierarchical K-NN classifier. Biomedical Imaging (ISBI), 2012 9th IEEE International Symposium on, 2012. IEEE, 358-361.

BALLERINI, L., FISHER, R. B., ALDRIDGE, B. & REES, J. 2013. A color and texture based hierarchical K-NN approach to the classification of non-melanoma skin lesions. *Color Medical Image Analysis.* Springer.

BARATA, C., CELEBI, M. E. & MARQUES, J. S. 2015. Improving dermoscopy image classification using color constancy. *IEEE journal of biomedical and health informatics,* 19**,** 1146-1152.

BARATA, C., MARQUES, J. S. & ROZEIRA, J. 2012. A system for the detection of pigment network in dermoscopy images using directional filters. *IEEE transactions on biomedical engineering,* 59**,** 2744-2754.

BARATA, C., MARQUES, J. S. & ROZEIRA, J. Evaluation of color based keypoints and features for the classification of melanomas using the bag-of-features model. International Symposium on Visual Computing, 2013. Springer, 40-49.

BARATA, C., RUELA, M., FRANCISCO, M., MENDONÇA, T. & MARQUES, J. S. 2014. Two systems for the detection of melanomas in dermoscopy images using texture and color features. *IEEE Systems Journal,* 8**,** 965-979.

BARHOUMI, W. & BAÂZAOUI, A. 2014. Pigment network detection in dermatoscopic images for melanoma diagnosis. *IRBM,* 35**,** 128-138.

BEHESHTI, Z. & SHAMSUDDIN, S. M. 2015. Non-parametric particle swarm optimization for global optimization. *Applied Soft Computing,* 28**,** 345-359.

BEHESHTI, Z. & SHAMSUDDIN, S. M. H. 2014. CAPSO: centripetal accelerated particle swarm optimization. *information Sciences,* 258**,** 54-79.

BERRY, L. 2016. Skin cancer. *Nursing standard (Royal College of Nursing (Great Britain): 1987),* 31**,** 15-15.

BEYAN, C. & FISHER, R. 2015. Classifying imbalanced data sets using similarity based hierarchical decomposition. *Pattern Recognition,* 48**,** 1653-1672.

BI, L., KIM, J., AHN, E., FENG, D. & FULHAM, M. Automatic melanoma detection via multi-scale lesion-biased representation and joint reverse classification. Biomedical Imaging (ISBI), 2016 IEEE 13th International Symposium on, 2016. IEEE, 1055-1058.

BI, L., KIM, J., AHN, E., KUMAR, A., FULHAM, M. & FENG, D. 2017. Dermoscopic image segmentation via multistage fully convolutional networks. *IEEE Transactions on Biomedical Engineering,* 64**,** 2065-2074.

BRADY, M. S., OLIVERIA, S. A., CHRISTOS, P. J., BERWICK, M., COIT, D. G., KATZ, J. & HALPERN, A. C. 2000. Patterns of detection in patients with cutaneous melanoma. *Cancer,* 89**,** 342-347.

BRASH, D. E., RUDOLPH, J. A., SIMON, J. A., LIN, A., MCKENNA, G. J., BADEN, H. P., HALPERIN, A. J. & PONTEN, J. 1991. A role for sunlight in skin cancer: UV-induced p53 mutations in squamous cell carcinoma. *Proceedings of the National Academy of Sciences,* 88**,** 10124-10128.

BRAY F, F. J., SOERJOMATARAM I, SIEGEL RL, TORRE LA, JEMAL A. 2018. *Global Cancer Statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries*. CA: A Cancer Journal for Clinicians.

BROWNE, M., GHIDARY, S. S. & MAYER, N. M. 2008. Convolutional neural networks for image processing with applications in mobile robotics. *Speech, Audio, Image and Biomedical Signal Processing using Neural Networks.* Springer.

BURGISS, S. G., JULIUS, C. E., WATSON, H. W., HAYNES, B. K., BUONOCORE, E. & SMITH, G. T. 1997. Telemedicine for dermatology care in rural patients. *Telemedicine Journal,* 3**,** 227-233.

CAMPBELL, D. 2014. Almost half of cancer patients diagnosed too late. *The Guardian*, 22 Sep 2014.

CANCER, I. A. F. R. O., W, S. B. & P, W. C. 2014. *World Cancer Report 2014 (PDF)*, International Agency for Research on Cancer (I A R C) (UN).

CARE, H. 2018. What is UVA and UVB? Available from: https://hazecare.com/blogs/news/what-is-uva-and-uvb.

CASCINELLI, N., FERRARIO, M., BUFALINO, R., ZURRIDA, S., GALIMBERTI, V., MASCHERONI, L., BARTOLI, C. & CLEMENTE, C. 1992. Results obtained by using a computerized image analysis system designed as an aid to diagnosis of cutaneous melanoma. *Melanoma Research,* 2**,** 163-170.

CHANDRAMOHAN, P. 2013. Medical education in India at crossroads: Issues and solutions. *Archives of Medicine and Health Sciences,* 1**,** 80.

CHANG, W.-D. 2015. A modified particle swarm optimization with multiple subpopulations for multimodal function optimization problems. *Applied Soft Computing,* 33**,** 170-182.

CHANG, W.-Y., HUANG, A., YANG, C.-Y., LEE, C.-H., CHEN, Y.-C., WU, T.-Y. & CHEN, G.-S. 2013. Computer-aided diagnosis of skin lesions using conventional digital photography: a reliability and feasibility study. *PloS one,* 8**,** e76212.

CHEN, Q., CHEN, Y. & JIANG, W. 2016. Genetic particle swarm optimization–based feature selection for very-high-resolution remotely sensed imagery object change detection. *Sensors,* 16**,** 1204.

CHEN, S. T., LI, X. & HAN, J. 2018. Personal history of non‐melanoma skin cancer diagnosis and death from melanoma in women. *International journal of cancer,* 142**,** 1536-1541.

CHEN, W.-N., ZHANG, J., LIN, Y., CHEN, N., ZHAN, Z.-H., CHUNG, H. S.-H., LI, Y. & SHI, Y.-H. 2013. Particle swarm optimization with an aging leader and challengers. *IEEE Transactions on Evolutionary Computation,* 17**,** 241-258.

CHEN, X., TIANFIELD, H., MEI, C., DU, W. & LIU, G. 2017. Biogeography-based learning particle swarm optimization. *Soft Computing,* 21**,** 7519-7541.

CHIEM, A., AL-JUMAILY, A. & KHUSHABA, R. N. A novel hybrid system for skin lesion detection. Intelligent Sensors, Sensor Networks and Information, 2007. ISSNIP 2007. 3rd International Conference on, 2007. IEEE, 567-572.

CHUANG, L.-Y., YANG, C.-H. & LI, J.-C. 2011a. Chaotic maps based on binary particle swarm optimization for feature selection. *Applied Soft Computing,* 11**,** 239-248.

CHUANG, L.-Y., YANG, C.-S., WU, K.-C. & YANG, C.-H. 2011b. Gene selection and classification using Taguchi chaotic binary particle swarm optimization. *Expert Systems with Applications,* 38**,** 13367-13377.

CLARIDGE, E., HALL, P., KEEFE, M. & ALLEN, J. 1992. Shape analysis for classification of malignant melanoma. *Journal of biomedical engineering,* 14**,** 229-234.

CLAUSI, D. A. & ZHAO, Y. 2002. Rapid extraction of image texture by co-occurrence using a hybrid data structure. *Computers & Geosciences,* 28**,** 763-774.

COTTON, S. & CLARIDGE, E. Developing a predictive model of human skin coloring. Medical Imaging 1996, 1996. International Society for Optics and Photonics, 814-825.

COURDAVAULT, S., BAUDOUIN, C., CHARVERON, M., CANGUILHEM, B., FAVIER, A., CADET, J. & DOUKI, T. 2005. Repair of the three main types of bipyrimidine DNA photoproducts in human keratinocytes exposed to UVB and UVA radiations. *DNA repair,* 4**,** 836-844.

DASH, J., DAM, B. & SWAIN, R. 2017. Optimal design of linear phase multi-band stop filters using improved cuckoo search particle swarm optimization. *Applied Soft Computing,* 52**,** 435-445.

DE GRUIJL, F. R., VAN KRANEN, H. J. & MULLENDERS, L. H. 2001. UV-induced DNA damage, repair, mutations and oncogenic pathways in skin cancer. *Journal of Photochemistry and Photobiology B: Biology,* 63**,** 19-27.

DEO, M. G. 2013. Doctor population ratio for India-The reality. *The Indian journal of medical research,* 137**,** 632.

DERMNET 2016. Dermnet Skin Desease Atlas. *In:* DERMNET (ed.).

DERMNET 2017. Dermnet Skin Disease Image Atlas.

DERRAC, J., GARCÍA, S., MOLINA, D. & HERRERA, F. 2011. A practical tutorial on the use of nonparametric statistical tests as a methodology for comparing evolutionary and swarm intelligence algorithms. *Swarm and Evolutionary Computation,* 1**,** 3-18.

DIEPGEN, T. & MAHLER, V. 2002. The epidemiology of skin cancer. *British Journal of Dermatology,* 146**,** 1-6.

DIFFEY, B. & FARR, P. 1991. Tanning with UVB or UVA: an appraisal of risks. *Journal of Photochemistry and Photobiology B: Biology,* 8**,** 219.

DOERR, B., HANSEN, N., IGEL, C. & THIELE, L. Theory of evolutionary algorithms (Dagstuhl seminar 15211). Dagstuhl Reports, 2016. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik.

DOUKAS, C., STAGKOPOULOS, P., KIRANOUDIS, C. T. & MAGLOGIANNIS, I. Automated skin lesion assessment using mobile technologies and cloud platforms. Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE, 2012. IEEE, 2444-2447.

DURO, D. C., FRANKLIN, S. E. & DUBÉ, M. G. 2012. A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery. *Remote Sensing of Environment,* 118**,** 259-272.

ELTAYEF, K., LI, Y. & LIU, X. Detection of pigment networks in dermoscopy images. Journal of Physics: Conference Series, 2017. IOP Publishing, 012033.

ESTEVA, A., KUPREL, B., NOVOA, R. A., KO, J., SWETTER, S. M., BLAU, H. M. & THRUN, S. 2017. Dermatologist-level classification of skin cancer with deep neural networks. *Nature,* 542**,** 115.

EXPOSED, B. L. 2018. Available: http://www.bluelightexposed.com [Accessed].

FENG, H., WU, P. & LEGER, M. 2016. Exploring the industry-dermatologist financial relationship: Insight from the open payment data. *JAMA dermatology,* 152**,** 1307-1313.

FERLAY, J., SHIN, H. R., BRAY, F., FORMAN, D., MATHERS, C. & PARKIN, D. M. 2010. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *International journal of cancer,* 127**,** 2893-2917.

FRANKE, K.-H., GAßMANN, F. & GERGS, R. Früherkennung von Hautkrebs durch Farbbildverarbeitung. Mustererkennung 1993, Mustererkennung im Dienste der Gesundheit, 15. DAGM-Symposium, 1993. Springer-Verlag, 407-414.

FRIEDBERG, E. C., WALKER, G. C., SIEDE, W. & WOOD, R. D. 2005. *DNA repair and mutagenesis,* American Society for Microbiology Press.

FRIEDMAN, R. & RIGEL, D. 1985. The clinical features of malignant melanoma. *Dermatologic clinics,* 3**,** 271-283.

GARG, N., SHARMA, V. & KAUR, P. 2018. Melanoma Skin Cancer Detection Using Image Processing. *Sensors and Image Processing.* Springer.

GEEM, Z. W., KIM, J. H. & LOGANATHAN, G. 2001. A new heuristic optimization algorithm: harmony search. *simulation,* 76**,** 60-68.

GILMORE, S., HOFMANN‐WELLENHOF, R. & SOYER, H. P. 2010. A support vector machine for decision support in melanoma recognition. *Experimental dermatology,* 19**,** 830-835.

GLAISTER, J., AMELARD, R., WONG, A. & CLAUSI, D. A. 2013. MSIM: Multistage illumination modeling of dermatological photographs for illumination-corrected skin lesion analysis. *IEEE Transactions on Biomedical Engineering,* 60**,** 1873-1883.

GLAISTER, J., WONG, A. & CLAUSI, D. A. 2014. Segmentation of skin lesions from digital images using joint statistical texture distinctiveness. *IEEE transactions on biomedical engineering,* 61**,** 1220-1230.

GLAZER, A. M., RIGEL, D. S., WINKELMANN, R. R. & FARBERG, A. S. 2017. Clinical Diagnosis of Skin Cancer: Enhancing Inspection and Early Recognition. *Dermatologic clinics,* 35**,** 409-416.

GOLA ISASI, A., GARCÍA ZAPIRAIN, B. & MÉNDEZ ZORRILLA, A. 2011. Melanomas non-invasive diagnosis application based on the ABCD rule and pattern recognition image processing algorithms. *Computers in Biology and Medicine,* 41**,** 742-755.

GOU, J., LEI, Y.-X., GUO, W.-P., WANG, C., CAI, Y.-Q. & LUO, W. 2017. A novel improved particle swarm optimization algorithm based on individual difference evolution. *Applied Soft Computing,* 57**,** 468-481.

GREEN, A., MARTIN, N., PFITZNER, J., O'ROURKE, M. & KNIGHT, N. 1994. Computer image analysis in the diagnosis of melanoma. *Journal of the American Academy of Dermatology,* 31**,** 958-964.

GRIFFITHS, C., BARKER, J., BLEIKER, T., CHALMERS, R. & CREAMER, D. 2016. *Rook's textbook of dermatology,* John Wiley & Sons.

GUTMAN, D., CODELLA, N. C., CELEBI, E., HELBA, B., MARCHETTI, M., MISHRA, N. & HALPERN, A. 2016. Skin lesion analysis toward melanoma detection: A challenge at the international symposium on biomedical imaging (ISBI) 2016, hosted by the international skin imaging collaboration (ISIC). *arXiv preprint arXiv:1605.01397*.

HABIF, T. P. 2015. *Clinical Dermatology E-Book,* Elsevier Health Sciences.

HAKLI, H. & UGUZ, H. 2013. Levy flight distribution for scout bee in artificial bee colony algorithm. *Lecture Notes on Software Engineering,* 1**,** 254.

HALL, P., CLARIDGE, E. & SMITH, J. M. 1995. Computer screening for early detection of melanoma—is there a future? *British Journal of Dermatology,* 132**,** 325-338.

HARALICK, R. M. 1979. Statistical and structural approaches to texture. *Proceedings of the IEEE,* 67**,** 786-804.

HEARST, M. A., DUMAIS, S. T., OSUNA, E., PLATT, J. & SCHOLKOPF, B. 1998. Support vector machines. *IEEE Intelligent Systems and their applications,* 13**,** 18-28.

HECKEL, L., FENNELL, K. M., MOHEBBI, M., BYRNES, M. & LIVINGSTON, P. M. 2017. Demographic characteristics, call details and psychosocial support needs of the family/friends of someone diagnosed with cancer who access Australian

Cancer Council telephone information and support services. *European Journal of Oncology Nursing,* 28**,** 86-91.

HOFFMANN, K., GAMBICHLER, T., RICK, A., KREUTZ, M., ANSCHUETZ, M., GRÜNENDICK, T., ORLIKOV, A., GEHLEN, S., PEROTTI, R. & ANDREASSI, L. 2003. Diagnostic and neural analysis of skin cancer (DANAOS). A multicentre study for collection and computer‐aided analysis of data from pigmented skin lesions using digital dermoscopy. *British Journal of Dermatology,* 149**,** 801-809.

HOLICK, M. F., BINKLEY, N. C., BISCHOFF-FERRARI, H. A., GORDON, C. M., HANLEY, D. A., HEANEY, R. P., MURAD, M. H. & WEAVER, C. M. 2011. Evaluation, treatment, and prevention of vitamin D deficiency: an Endocrine Society clinical practice guideline. *The Journal of Clinical Endocrinology & Metabolism,* 96**,** 1911-1930.

HOLLAND, J. H. 1975. Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence. University of Michigan Press Ann Arbor.

HOLLAND, J. H. 1992. Genetic algorithms. *Scientific american,* 267**,** 66-73.

HU, Z., BAO, Y. & XIONG, T. 2014. Comprehensive learning particle swarm optimization based memetic algorithm for model selection in short-term load forecasting using support vector regression. *Applied Soft Computing,* 25**,** 15-25.

JAIN, A. K., MAO, J. & MOHIUDDIN, K. M. 1996. Artificial neural networks: A tutorial. *Computer,* 29**,** 31-44.

JAMIL, U., KHALID, S., AKRAM, M. U., AHMAD, A. & JABBAR, S. 2018. Melanocytic and nevus lesion detection from diseased dermoscopic images using fuzzy and wavelet techniques. *Soft Computing,* 22**,** 1577-1593.

JEMAL, A., BRAY, F., CENTER, M. M., FERLAY, J., WARD, E. & FORMAN, D. 2011. Global cancer statistics. *CA: a cancer journal for clinicians,* 61**,** 69-90.

JERANT, A. F., JOHNSON, J. T., DEMASTES SHERIDAN, C. & CAFFREY, T. J. 2000. Early detection and treatment of skin cancer. *American family physician,* 62.

JORDEHI, A. R. 2015. Enhanced leader PSO (ELPSO): a new PSO variant for solving global optimisation problems. *Applied Soft Computing,* 26**,** 401-417.

KARABOGA, D. & BASTURK, B. 2007. A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm. *Journal of global optimization,* 39**,** 459-471.

KARRAN, P. & BREM, R. 2016. Protein oxidation, UVA and human DNA repair. *DNA repair,* 44**,** 178-185.

KASMI, R. & MOKRANI, K. 2016. Classification of malignant melanoma and benign skin lesions: implementation of automatic ABCD rule. *IET Image Processing,* 10**,** 448-455.

KATAPADI, A. B., CELEBI, M. E., TROTTER, S. C. & GURCAN, M. N. 2017. Evolving strategies for the development and evaluation of a computerised melanoma image analysis system. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization***,** 1-8.

KAVEH, A. & TALATAHARI, S. 2010. A novel heuristic optimization method: charged system search. *Acta Mechanica,* 213**,** 267-289.

KENNEDY, J. 2011. Particle swarm optimization. *Encyclopedia of machine learning.* Springer.

KHAN, A. Q., TRAVERS, J. B. & KEMP, M. G. 2018. Roles of UVA radiation and DNA damage responses in melanoma pathogenesis. *Environmental and molecular mutagenesis.*

KIANI, K. & SHARAFAT, A. R. 2011. E-shaver: An improved DullRazor® for digitally removing dark and light-colored hairs in dermoscopic images. *Computers in biology and medicine,* 41**,** 139-145.

KIMESWENGER, S., DINGELMAIER‐HOVORKA, R., FOEDINGER, D. & JANTSCHITSCH, C. 2018. UVA1 impairs the repair of UVB‐induced DNA damage in normal human melanocytes. *Experimental dermatology,* 27**,** 276-279.

KıRAN, M. S. & FıNDıK, O. 2015. A directed artificial bee colony algorithm. *Applied Soft Computing,* 26**,** 454-462.

KOZA, J. R. 1992. *Genetic programming: on the programming of computers by means of natural selection*, MIT press.

KRISSHNA, N. A., DEEPAK, V. K., MANIKANTAN, K. & RAMACHANDRAN, S. 2014. Face recognition using transform domain feature extraction and PSO-based feature selection. *Applied Soft Computing,* 22**,** 141-161.

KROHLING, R. A. & MENDEL, E. Bare bones particle swarm optimization with Gaussian or Cauchy jumps. Evolutionary Computation, 2009. CEC'09. IEEE Congress on, 2009. IEEE, 3285-3291.

KRUK, M., ŚWIDERSKI, B., OSOWSKI, S., KUREK, J., SŁOWIŃSKA, M. & WALECKA, I. 2015. Melanoma recognition using extended set of descriptors and classifiers. *EURASIP Journal on Image and Video Processing*, 2015**,** 43.

LE CLAIR, M. Z. & COCKBURN, M. G. 2016. Tanning bed use and melanoma: Establishing risk and improving prevention interventions. *Preventive medicine reports,* 3**,** 139-144.

LEE, H.-C. 1994. Skin cancer diagnosis using hierarchical neural networks and fuzzy logic.

LEE, J. H., KWON, H. S., JUNG, H. M., KIM, G. M. & BAE, J. M. 2018. Wood's lamp–induced fluorescence of milia. *Journal of the American Academy of Dermatology,* 78**,** e99-e100.

LÉVY, P. & BOREL, É. 1954. Théorie de l'addition des variables aléatoires.

LI, N.-J., WANG, W.-J. & HSU, C.-C. J. 2015. Hybrid particle swarm optimization incorporating fuzzy reasoning and weighted particle. *Neurocomputing,* 167**,** 488-501.

LI, Y. & SHEN, L. 2018. Skin lesion analysis towards melanoma detection using deep learning network. *Sensors,* 18**,** 556.

LICHMAN, M. 2013. UCI Machine Learning Repository. *In:* IRVINE, C. U. O. C., SCHOOL OF INFORMATION AND COMPUTER SCIENCE (ed.).

LIM, W. H. & ISA, N. A. M. 2015. Adaptive division of labor particle swarm optimization. *Expert Systems with Applications,* 42**,** 5887-5903.

LIMONE, B. & MEADOWS, L. 2017. The Wood's Light's Diagnostic Use in Dermatology. *Journal of the Dermatology Nurses' Association,* 9**,** 211-215.

LINOS, E., SWETTER, S. M., COCKBURN, M. G., COLDITZ, G. A. & CLARKE, C. A. 2009. Increasing burden of melanoma in the United States. *Journal of Investigative Dermatology,* 129**,** 1666-1674.

LYNN, N. & SUGANTHAN, P. N. 2015. Heterogeneous comprehensive learning particle swarm optimization with enhanced exploration and exploitation. *Swarm and Evolutionary Computation,* 24**,** 11-24.

LYNN, N. & SUGANTHAN, P. N. 2017. Ensemble particle swarm optimizer. *Applied Soft Computing,* 55**,** 533-548.

MAHMOODABADI, M. J., MOTTAGHI, Z. S. & BAGHERI, A. 2014. HEPSO: high exploration particle swarm optimization. *Information Sciences,* 273**,** 101-111.

MAHMOUD, H., ABDEL-NASSER, M. & OMER, O. A. Computer aided diagnosis system for skin lesions detection using texture analysis methods. Innovative Trends in Computer Engineering (ITCE), 2018 International Conference on, 2018. IEEE, 140-144.

MARK HOPKINS, E. R., GEORGE FORMAN, JAAP SUERMONDT, HEWLETT-PACKARD LABS 1998. Spambase.

MARQUES, J. S., BARATA, C. & MENDONCA, T. On the role of texture and color in the classification of dermoscopy images. Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE, 2012. IEEE, 4402-4405.

MCGUIRE, S. T., SECREST, A. M., ANDRULONIS, R. & FERRIS, L. K. 2011. Surveillance of patients for early detection of melanoma: patterns in dermatologist vs patient discovery. *Archives of dermatology,* 147**,** 673-678.

MENDONCŸA, T., FERREIRA, P., MARQUES, J., MARCŸAL, A. & ROZEIRA, J. J. P. I. P. O. P. 2013. A dermoscopic image database for research and benchmarking. 2.

MENZIES, S. W., CROTTY, K. A., INGVAR, C. & MCCARTHY, W. H. 2003. *An atlas of surface microscopy of pigmented skin lesions: dermoscopy*, McGraw-Hill Sydney, Australia.

MILAN, S. 2017. Skin Biopsies: Punch, Shave and Excisional. *Urgent Procedures in Medical Practice***,** 31.

MILTON, M. A. A. J. A. P. A. 2019. Automated Skin Lesion Classification Using Ensemble of Deep Neural Networks in ISIC 2018: Skin Lesion Analysis Towards Melanoma Detection Challenge.

MIRJALILI, S. 2015. Moth-flame optimization algorithm: A novel nature-inspired heuristic paradigm. *Knowledge-Based Systems,* 89**,** 228-249.

MIRJALILI, S. 2016. Dragonfly algorithm: a new meta-heuristic optimization technique for solving single-objective, discrete, and multi-objective problems. *Neural Computing and Applications,* 27**,** 1053-1073.

MIRJALILI, S., LEWIS, A. & SADIQ, A. S. 2014a. Autonomous particles groups for particle swarm optimization. *Arabian Journal for Science and Engineering,* 39**,** 4683-4697.

MIRJALILI, S., MIRJALILI, S. M. & LEWIS, A. 2014b. Grey wolf optimizer. *Advances in Engineering Software,* 69**,** 46-61.

MISTRY, K., ZHANG, L., NEOH, S. C., LIM, C. P. & FIELDING, B. 2017. A micro-GA embedded PSO feature selection approach to intelligent facial emotion recognition. *IEEE transactions on cybernetics,* 47**,** 1496-1509.

MOGHADDAM, F. F., MOGHADDAM, R. F. & CHERIET, M. 2012. Curved space optimization: a random search based on general relativity theory. *arXiv preprint arXiv:1208.2214.*

MONISHA, M., SURESH, A., BAPU, B. R. T. & RASHMI, M. R. 2018. Classification of malignant melanoma and benign skin lesion by using back propagation neural network and ABCD rule. *Cluster Computing*.

MORRIS, S., COX, B. & BOSANQUET, N. 2009. Cost of skin cancer in England. *The European Journal of Health Economics,* 10**,** 267-273.

NACHBAR, F., STOLZ, W., MERKLE, T., COGNETTA, A. B., VOGT, T., LANDTHALER, M., BILEK, P., BRAUN-FALCO, O. & PLEWIG, G. 1994. The ABCD rule of dermatoscopy. *Journal of the American Academy of Dermatology,* 30**,** 551-559.

NASIR, M., DAS, S., MAITY, D., SENGUPTA, S., HALDER, U. & SUGANTHAN, P. N. 2012. A dynamic neighborhood learning based particle swarm optimizer for global numerical optimization. *Information Sciences,* 209**,** 16-36.

NAVARRO, F., ESCUDERO-VINOLO, M. & BESCOS, J. 2018. Accurate segmentation and registration of skin lesion images to evaluate lesion change. *IEEE Journal of Biomedical and Health Informatics***,** 1-1.

NEOH, S. C., SRISUKKHAM, W., ZHANG, L., TODRYK, S., GREYSTOKE, B., LIM, C. P., HOSSAIN, M. A. & ASLAM, N. 2015. An intelligent decision support system for leukaemia diagnosis using microscopic blood images. *Scientific reports,* 5**,** 14938.

NEVADA, C. C. C. O. 2017. UVA vs. UVB Rays: What the Difference Is and Why You Need to Protect Your Skin Against Both. Comprehensive Cancer Centers of Nevada.

NICHOLS, J. A. & KATIYAR, S. K. 2010. Skin photoprotection by natural polyphenols: anti-inflammatory, antioxidant and DNA repair mechanisms. *Archives of dermatological research,* 302**,** 71-83.

OUYANG, H.-B., GAO, L.-Q., LI, S. & KONG, X.-Y. 2017. Improved global-best-guided particle swarm optimization with learning operation for global optimization problems. *Applied Soft Computing,* 52**,** 987-1008.

PATTERSON, J. W. 2014. *Weedon's Skin Pathology E-Book*, Elsevier Health Sciences.

PENNISI, A., BLOISI, D. D., NARDI, D., GIAMPETRUZZI, A. R., MONDINO, C. & FACCHIANO, A. 2016. Skin lesion image segmentation using Delaunay Triangulation for melanoma detection. *Computerized Medical Imaging and Graphics,* 52**,** 89-103.

PETMANSON, T. 2011. Ant Colony Optimisation (ACO).

PFEIFER, G. P., YOU, Y.-H. & BESARATINIA, A. 2005. Mutations induced by ultraviolet light. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis,* 571**,** 19-31.

PICKETT, H. S. A. P. B. F. S. L. 2011. Am Fam Physician.

QIN, Q., CHENG, S., ZHANG, Q., WEI, Y. & SHI, Y. 2015. Multiple strategies based orthogonal design particle swarm optimizer for numerical optimization. *Computers & Operations Research,* 60**,** 91-110.

QUEEN, L. 2017. Skin Cancer: Causes, Prevention, and Treatment.

RAMLAKHAN, K. & SHANG, Y. A mobile automated skin lesion classification system. Tools with Artificial Intelligence (ICTAI), 2011 23rd IEEE International Conference on, 2011. IEEE, 138-141.

RASHEDI, E., NEZAMABADI-POUR, H. & SARYAZDI, S. 2009. GSA: A Gravitational Search Algorithm. *Information Sciences,* 179**,** 2232-2248.

RECHENBERG, I. 1989. Evolution strategy: Nature's way of optimization. *Optimization: Methods and applications, possibilities and limitations.* Springer.

REYNOLDS, R. G. 1999. Cultural algorithms: theory and applications. *In:* DAVID, C., MARCO, D., FRED, G., DIPANKAR, D., PABLO, M., RICCARDO, P. & KENNETH, V. P. (eds.) *New ideas in optimization.* McGraw-Hill Ltd., UK.

RIAZ, F., HASSAN, A., JAVED, M. Y. & COIMBRA, M. T. Detecting melanoma in dermoscopy images using scale adaptive local binary patterns. Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE, 2014. IEEE, 6758-6761.

ROKACH, L. 2010. Ensemble-based classifiers. *Artificial Intelligence Review,* 33**,** 1-39.

SÁEZ, A., ACHA, B. & SERRANO, C. 2014. Pattern analysis in dermoscopic images. *Computer Vision Techniques for the Diagnosis of Skin Cancer.* Springer.

SAHINER, B., CHAN, H.-P., PETRICK, N., WEI, D., HELVIE, M. A., ADLER, D. D. & GOODSITT, M. M. 1996. Classification of mass and normal breast tissue: a convolution neural network classifier with spatial domain and texture images. *IEEE transactions on Medical Imaging,* 15**,** 598-610.

SATHEESHA, T., SATYANARAYANA, D., PRASAD, M. G. & DHRUVE, K. D. 2017. Melanoma Is Skin Deep: A 3D Reconstruction Technique for Computerized Dermoscopic Skin Lesion Classification. *IEEE Journal of Translational Engineering in Health and Medicine,* 5**,** 1-17.

SCHMID, P. 1999. Segmentation of digitized dermatoscopic images by two-dimensional color clustering. *IEEE Transactions on Medical Imaging,* 18**,** 164-171.

SEIDENARI, S., BURRONI, M., DELL'EVA, G., PEPE, P. & BELLETTI, B. 1995. Computerized evaluation of pigmented skin lesion images recorded by a videomicroscope: comparison between polarizing mode observation and oil/slide mode observation. *Skin Research and Technology,* 1**,** 187-191.

SHI, Y. & EBERHART, R. A modified particle swarm optimizer. Evolutionary Computation Proceedings, 1998. IEEE World Congress on Computational Intelligence., The 1998 IEEE International Conference on, 1998. IEEE, 69-73.

SHIMIZU, K., IYATOMI, H., CELEBI, M. E., NORTON, K.-A. & TANAKA, M. 2015. Four-class classification of skin lesions with task decomposition strategy. *IEEE Transactions on Biomedical Engineering,* 62**,** 274-283.

SIEGEL, R. L., MILLER, K. D. & JEMAL, A. 2016. Cancer statistics, 2016. *CA: a cancer journal for clinicians,* 66**,** 7-30.

SIEGEL, R. L., MILLER, K. D. & JEMAL, A. 2018. Cancer statistics, 2018. *CA: a cancer journal for clinicians,* 68**,** 7-30.

SILVEIRA, M., NASCIMENTO, J. C., MARQUES, J. S., MARÇAL, A. R., MENDONÇA, T., YAMAUCHI, S., MAEDA, J. & ROZEIRA, J. 2009. Comparison of segmentation methods for melanoma diagnosis in dermoscopy images. *IEEE Journal of Selected Topics in Signal Processing,* 3**,** 35-45.

SIVARAJ, R. & RAVICHANDRAN, T. 2011. A review of selection methods in genetic algorithm. *International journal of engineering science and technology,* 3**,** 3792-3797.

SMAOUI, N. & BESSASSI, S. 2013. A developed system for melanoma diagnosis. *International Journal of Computer Vision and Signal Processing,* 3**,** 10-17.

SNEDECOR, G. W. & COCHRAN, W. G. 1967. Statistical methods 6th edition. *The Iowa State University*.

SOUMYA, R., NEETHU, S., NIJU, T., RENJINI, A. & ANEESH, R. Advanced earlier melanoma detection algorithm using colour correlogram. Communication Systems and Networks (ComNet), International Conference on, 2016. IEEE, 190-194.

SRISUKKHAM, W., ZHANG, L., NEOH, S. C., TODRYK, S. & LIM, C. P. 2017. Intelligent leukaemia diagnosis with bare-bones PSO based feature optimization. *Applied Soft Computing*, 56**,** 405-419.

STOECKER, W. V., LI, W. W. & MOSS, R. H. 1992. Automatic detection of asymmetry in skin tumors. *Computerized Medical Imaging and Graphics*, 16**,** 191-197.

STORN, R. & PRICE, K. 1997. Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces. *Journal of global optimization*, 11**,** 341-359.

SUTO, J., ONIGA, S. & SITAR, P. P. Comparison of wrapper and filter feature selection algorithms on human activity recognition. Computers Communications and Control (ICCCC), 2016 6th International Conference on, 2016. IEEE, 124-129.

TAHERKHANI, M. & SAFABAKHSH, R. 2016. A novel stability-based adaptive inertia weight for particle swarm optimization. *Applied Soft Computing*, 38**,** 281-295.

TALBI, E.-G., JOURDAN, L., GARCIA-NIETO, J. & ALBA, E. Comparison of population based metaheuristics for feature selection: Application to microarray data classification. Computer Systems and Applications, 2008. AICCSA 2008. IEEE/ACS International Conference on, 2008. IEEE, 45-52.

TAN, T. Y., ZHANG, L. & JIANG, M. An intelligent decision support system for skin cancer detection from dermoscopic images. Natural Computation, Fuzzy Systems

and Knowledge Discovery (ICNC-FSKD), 2016 12th International Conference on, 2016. IEEE, 2194-2199.

TAN, T. Y., ZHANG, L., NEOH, S. C. & LIM, C. P. J. K.-B. S. 2018. Intelligent skin cancer detection using enhanced particle swarm optimization. 158**,** 118-135.

TANG, K., LI, Z., LUO, L. & LIU, B. 2015. Multi-strategy adaptive particle swarm optimization for numerical optimization. *Engineering Applications of Artificial Intelligence,* 37**,** 9-19.

TAUFIQ, M. A., HAMEED, N., ANJUM, A. & HAMEED, F. 2017. m-Skin Doctor: A Mobile Enabled System for Early Melanoma Skin Cancer Detection Using Support Vector Machine. *eHealth 360°*. Springer.

THARWAT, A., HASSANIEN, A. E. & ELNAGHI, B. E. 2017. A ba-based algorithm for parameter optimization of support vector machine. *Pattern Recognition Letters,* 93**,** 13-22.

TITUS, L., CLOUGH‐GORR, K., MACKENZIE, T., PERRY, A., SPENCER, S., WEISS, J., ABRAHAMS‐GESSEL, S. & ERNSTOFF, M. 2013. Recent skin self‐examination and doctor visits in relation to melanoma risk and tumour depth. *British journal of dermatology,* 168**,** 571-576.

TRAVEL, U. C. 2015. Melanoma: Signs and Prevention.

UDREA, A. & LUPU, C. Real-time acquisition of quality verified nonstandardized color images for skin lesions risk assessment—A preliminary study. System Theory, Control and Computing (ICSTCC), 2014 18th International Conference, 2014. IEEE, 199-204.

ÜNAL, M., AK, A., TOPUZ, V. & ERDAL, H. 2013. Ant Colony Optimization (ACO). *Optimization of PID Controllers Using Ant Colony and Genetic Algorithms.* Springer.

UNLER, A. & MURAT, A. 2010. A discrete particle swarm optimization method for feature selection in binary classification problems. *European Journal of Operational Research,* 206**,** 528-539.

VALENTINE, M. L. 2017. Skin Cancer. *The Well-Woman Visit***,** 204.

VASCONCELOS, M. J. M., ROSADO, L. & FERREIRA, M. A new color assessment methodology using cluster-based features for skin lesion analysis. Information and Communication Technology, Electronics and Microelectronics (MIPRO), 2015 38th International Convention on, 2015. IEEE, 373-378.

WAHEED, Z., WAHEED, A., ZAFAR, M. & RIAZ, F. An efficient machine learning approach for the detection of melanoma using dermoscopic images. Communication, Computing and Digital Systems (C-CODE), International Conference on, 2017. IEEE, 316-319.

WANG, H., RAHNAMAYAN, S., SUN, H. & OMRAN, M. G. 2013. Gaussian bare-bones differential evolution. *Cybernetics, IEEE Transactions on,* 43**,** 634-647.

WEBB, A. R., KAZANTZIDIS, A., KIFT, R. C., FARRAR, M. D., WILKINSON, J. & RHODES, L. E. 2018. Colour Counts: Sunlight and skin type as drivers of vitamin D deficiency at UK latitudes. *Nutrients,* 10**,** 457.

WEBSTER, B. & BERNHARD, P. J. 2003. A local search optimization algorithm based on natural principles of gravitation.

WILCOXON, F. 1945. Individual Comparisons by Ranking Methods. *Biometrics Bulletin,* 1**,** 80-83.

WINSLOW, T. 2009. Skin Cancer, Clark Levels. National Cancer Institute.

XIE, F., FAN, H., LI, Y., JIANG, Z., MENG, R. & BOVIK, A. 2017. Melanoma classification on dermoscopy images using a neural network ensemble model. *IEEE transactions on medical imaging,* 36**,** 849-858.

XIE, L. & ZENG, J. 2010. The performance analysis of artificial physics optimization algorithm driven by different virtual forces. *ICIC Express Letters (ICIC-EL),* 4**,** 239-244.

XUE, B., ZHANG, M. & BROWNE, W. N. 2013. Particle swarm optimization for feature selection in classification: A multi-objective approach. *IEEE transactions on cybernetics,* 43**,** 1656-1671.

YAGERMAN, S. & STEVENSON, M. L. 2018. Procedures in the Diagnosis and Treatment of Skin Cancer. *A Practical Guide to Skin Cancer.* Springer.

YANG, L. & XU, Z. 2017. Feature extraction by PCA and diagnosis of breast tumors using SVM with DE-based parameter tuning. *International Journal of Machine Learning and Cybernetics***,** 1-11.

YANG, Q., TIAN, J. & SI, W. 2017. An improved particle swarm optimization based on difference equation analysis. *Journal of Difference Equations and Applications,* 23**,** 135-152.

YANG, X.-S. 2010a. Firefly algorithm, Levy flights and global optimization. *Research and development in intelligent systems XXVI***,** 209-218.

YANG, X.-S. 2010b. A new metaheuristic bat-inspired algorithm. *Nature inspired cooperative strategies for optimization (NICSO 2010)***,** 65-74.

YANG, X.-S. Flower pollination algorithm for global optimization. International Conference on Unconventional Computing and Natural Computation, 2012. Springer, 240-249.

YANG, X.-S. & DEB, S. Cuckoo search via Lévy flights. Nature & Biologically Inspired Computing, 2009. NaBIC 2009. World Congress on, 2009. IEEE, 210-214.

YANG, X.-S. & HOSSEIN GANDOMI, A. 2012. Bat algorithm: a novel approach for global engineering optimization. *Engineering Computations,* 29**,** 464-483.

YAO, X., LIU, Y. & LIN, G. 1999. Evolutionary programming made faster. *IEEE Transactions on Evolutionary computation,* 3**,** 82-102.

YOGESH, C., HARIHARAN, M., NGADIRAN, R., ADOM, A. H., YAACOB, S., BERKAI, C. & POLAT, K. 2017. A new hybrid PSO assisted biogeography-based optimization for emotion and stress recognition from speech signal. *Expert Systems with Applications,* 69**,** 149-158.

ZHAN, Z.-H., ZHANG, J., LI, Y. & SHI, Y.-H. 2011. Orthogonal learning particle swarm optimization. *IEEE transactions on evolutionary computation,* 15**,** 832-847.

ZHANG, L., JIANG, M., FARID, D. & HOSSAIN, M. A. 2013. Intelligent facial emotion recognition and semantic-based topic detection for a humanoid robot. *Expert Systems with Applications,* 40**,** 5160-5168.

ZHANG, L., MISTRY, K., LIM, C. P. & NEOH, S. C. 2018a. Feature selection using firefly optimization for classification and regression models. *Decision Support Systems,* 106**,** 64-85.

ZHANG, L., SRISUKKHAM, W., NEOH, S. C., LIM, C. P. & PANDIT, D. 2018b. Classifier ensemble reduction using a modified firefly algorithm: An empirical evaluation. *Expert Systems with Applications,* 93**,** 395-422.

ZHANG, Y., GONG, D., HU, Y. & ZHANG, W. 2015a. Feature selection algorithm based on bare bones particle swarm optimization. *Neurocomputing,* 148**,** 150-157.

ZHANG, Y., ZHANG, L. & HOSSAIN, M. A. 2015b. Adaptive 3D facial action intensity estimation and emotion recognition. *Expert Systems with Applications,* 42**,** 1446-1464.

ZHANG, Y., ZHANG, L., NEOH, S. C., MISTRY, K. & HOSSAIN, M. A. 2015c. Intelligent affect regression for bodily expressions using hybrid particle swarm optimization and adaptive ensembles. *Expert Systems with Applications,* 42**,** 8678-8697.